

Random MAX SAT, Random MAX CUT, and Their Phase Transitions

Don Coppersmith* David Gamarnik*

Mohammad Hajiaghayi† Gregory B. Sorkin*

4 March 2003

Abstract

Given a 2-SAT formula F consisting of n variables and $\lfloor cn \rfloor$ random clauses, what is the largest number of clauses $\max F$ satisfiable by a single assignment of the variables? We bound the answer away from the trivial bounds of $\frac{3}{4}cn$ and cn . We prove that for $c < 1$, the expected number of clauses satisfiable is $\lfloor cn \rfloor - \Theta(1/n)$; for large c , it is $(\frac{3}{4}c + \Theta(\sqrt{c}))n$; for $c = 1 + \varepsilon$, it is at least $(1 + \varepsilon - O(\varepsilon^3))n$ and at most $(1 + \varepsilon - \Omega(\varepsilon^3 / \ln \varepsilon))n$; and in the “scaling window” $c = 1 + \Theta(n^{-1/3})$, it is $cn - \Theta(1)$. In particular, just as the decision problem undergoes a phase transition, our optimization problem also undergoes a phase transition at the same critical value $c = 1$.

Nearly all of our results are established without reference to the analogous propositions for decision 2-SAT, and as a byproduct we reproduce many of those results, including much of what is known about the 2-SAT scaling window.

We consider “online” versions of MAX 2-SAT, and show that for one version, the obvious greedy algorithm is optimal.

We can extend only our simplest MAX 2-SAT results to MAX k -SAT, but we conjecture a “MAX k -SAT limiting function conjecture” analogous to the folklore satisfiability threshold conjecture, but open even for $k = 2$. Neither conjecture immediately implies the other, but it is natural to further conjecture a connection between them.

Finally, for random MAX CUT (the size of a maximum cut in a sparse random graph) we prove analogous results.

*Department of Mathematical Sciences, IBM T.J. Watson Research Center, Yorktown Heights NY 10598, USA. e-mail {copper,gamarnik,sorkin}@watson.ibm.com

†Department of Mathematics, M.I.T., Cambridge MA 02139, USA.
e-mail hajigha@math.mit.edu

Contents

1	Introduction	3
1.1	Outlook	3
1.2	Problem and motivations	4
2	Notation and model	6
3	Summary of results	6
4	Random MAX 2-SAT	9
4.1	Sub-critical MAX 2-SAT	9
4.2	High-density random MAX 2-SAT	12
4.3	Low-density random MAX 2-SAT	14
5	The MAX 2-SAT scaling window	17
5.1	Case $c = 1 + \lambda n^{-1/3}$, $\lambda \leq -1$	18
5.2	Case $c = 1 + \lambda n^{-1/3}$, $\lambda \geq 1$	21
5.2.1	Useful facts	22
5.2.2	Phase I	24
5.2.3	Phase II	27
5.2.4	Phases I, II and III	29
5.2.5	Remarks	30
6	Random MAX k-SAT and MAX CSP	30
6.1	Concentration and limits	30
6.2	High-density MAX k-SAT and MAX CSP	32
7	Online random MAX 2-SAT	34
8	Random MAX CUT	37
8.1	Motivation	37
8.2	MAX CUT	39
8.3	Results	40
8.4	Subcritical MAX CUT	41
8.5	High-density random MAX CUT	41
8.6	Low-density random MAX CUT	42
8.7	Scaling window	43
9	Conclusions and open problems	45

1 Introduction

In this paper, we consider random instances of MAX 2-SAT, MAX k -SAT, and MAX CUT. Just as random instances of the decision problem 2-SAT show a phase transition from almost-sure satisfiability to almost-sure unsatisfiability as the instance “density” increases above 1, so the maximization problem shows a transition at the same point, with the expected number of clauses *not* satisfied by an optimal solution quickly changing from $\Theta(1/n)$ to $\Theta(n)$. MAX CUT experiences a similar phase transition: as a random graph’s edge density crosses above $1/n$, the number of edges *not* cut in an optimal cut suddenly changes from $\Theta(1)$ to $\Theta(n)$.

Our methods are well established ones: the first-moment method for upper bounds; algorithmic analysis including the differential-equation method for lower bounds; and some more sophisticated arguments for the analysis of the scaling window. The interest of the work lies in the simplicity of the methods, and in the results. The questions we ask seem very natural, and the answers obtained for MAX 2-SAT and MAX CUT are happily neat, and fairly comprehensive.

A preliminary version of this paper appeared as [CGHS03].

1.1 Outlook

Beyond our particular results for MAX 2-SAT and MAX CUT, we hope to spark further work on phase transitions in random instances of other optimization problems, in particular of MAX CSPs (constraint satisfaction problems). Random instances of optimization problems have been studied extensively — some that come to mind are the travelling salesman problem, minimum spanning tree, minimum assignment, minimum bisection, minimum coloring, and maximum clique — but little has been said about *phase transitions* in such cases, and indeed many of the examples do not even have a natural parameter whose continuous variation could give rise to a phase transition.

Many problems, including all CSPs, have natural decision and optimization versions: one can ask whether a graph is k -colorable, or ask for the minimum number of colors it requires. We suggest that in a random setting, the optimization version is quite as interesting as the decision version. Furthermore, optimization problems may plausibly be easier to analyze than decision problems because the quantities of interest vary more smoothly. In fact, a recent triumph in the analysis of a decision problem, the characterization of the “scaling window” for 2-SAT, used as a smoothed quantity the size of the “spine” of a formula [BBC⁺01]. A way to view our MAX 2-SAT results

is that instead of taking the size of the spine as our “order parameter”, we take the size of a maximum satisfiable subformula. This seems comparably tractable (we reproduce the result of [BBC⁺01] incompletely, but more easily), and arguably more natural. Generally, when a decision problem has an optimization analog, the value of the optimum is both interesting in its own right, and, we suggest, an obvious candidate order parameter for studying the decision problem.

1.2 Problem and motivations

Let F be a k -SAT formula with n variables X_1, \dots, X_n . An “assignment” of these variables consists of setting each X_i to either 1 (True) or 0 (False); we may write an assignment as a vector $\vec{X} \in \{0, 1\}^n$. k -SAT is well understood. In particular, it is a canonical NP-hard problem to determine if a given formula F is satisfiable or not, except for $k = 2$ when this decision problem is solvable in essentially linear time.

Random instances of k -SAT have recently received wide attention. Let $\mathcal{F}(n, m)$ denote the set of all formulas with n variables and m clauses, where each clause is proper (consisting of k distinct variables, each of which may be complemented or not), and clauses may be repeated. Let $F \in \mathcal{F}$ be chosen uniformly at random; this is equivalent to choosing m clauses uniformly at random, with replacement, from the $2^k \binom{n}{k}$ possible clauses.

The model is generally parametrized as $F \in \mathcal{F}(n, cn)$ for various “densities” c , and the state of knowledge is summarized thus. The 2-SAT case is well understood: for $c < 1$, F is almost surely satisfiable (a.s. in the limit $n \rightarrow \infty$), and for $c > 1$, F is a.s. unsatisfiable [CR92, Goe96, FdV92]. Recently, the “scaling window” $c = 1 \pm \Theta(n^{-1/3})$ has also been analyzed [BBC⁺01]. For k -SAT, much less is known. For 3-SAT, for instance, it is known that for $c < 3.42$, F is a.s. satisfiable [KKL02] and for $c > 4.6$, F is a.s. unsatisfiable [JSV00]. It is only conjectured, though, that for $k = 3$ (and for all k) the situation is similar to that for $k = 2$.

Conjecture 1 (Satisfiability Threshold Conjecture) *For each k there exists a threshold density c_k , such that for any positive ε , for all $c < c_k - \varepsilon$, a random formula F is a.s. satisfiable, and for all $c > c_k + \varepsilon$, F is a.s. unsatisfiable.*

For large values of k , although the question of a threshold remains open, satisfiability and unsatisfiability density bounds are asymptotically equal, as shown by an analysis in [AM02] and refined in [AP03]. The closest result to the satisfiability conjecture is a theorem of Friedgut [Fri99] proving similar

thresholds, but leaving open the possibility that (for a given k), each n may have its own threshold, and that these may not converge to a limit.

Theorem 2 (Friedgut) *For each k there exists a threshold density function $c_k(n)$, such that for any positive ε , as $n \rightarrow \infty$, for all $c < c_k - \varepsilon$, a random formula F is a.s. satisfiable, and for all $c > c_k + \varepsilon$, F is a.s. unsatisfiable.*

Having briefly surveyed *random k -SAT*, let us similarly consider *max k -sat*. For a given formula F , let $F(\vec{X})$ be the number of clauses satisfied by \vec{X} . The problem MAX 2-SAT asks for $\max F \doteq \max_{\vec{X}} F(\vec{X})$, i.e., the maximum, over all assignments \vec{X} , of the size (number of clauses) of a maximum satisfiable subformula of F .

In the maximization setting, even 2-SAT is interesting. MAX 2-SAT is NP-hard to solve exactly, and it is even NP-hard to approximate $\max F$ to within a factor of $21/22$ [Hås97]. On the other hand, a $3/4$ -approximation is trivial: a random assignment satisfies an expected $3/4$ ths of the clauses, and a derandomized algorithm is simple (our algorithm used to prove the lower bound for Theorem 4 can serve). The best known approximation ratio achievable in polynomial time is 0.940 [LLZ02]. For arbitrary 3-SAT formulas F , in polynomial time, $\max F$ can be approximated to within a factor of $7/8$ [KZ97], but no better (unless P=NP) [Hås97].

Although both randomized and maximization versions of k -SAT are thus well studied, we are aware of no work on random MAX SAT, nor other random MAX or MIN constraint satisfaction problems (CSPs). These problems seem very natural, and answers to even the simplest questions are not obvious at first blush: For a random 2-SAT formula $F(n, cn)$ with $c > 1$, which is a.s. unsatisfiable, can we perhaps w.h.p. satisfy all but a single clause?

These questions have elegant answers; we will show for example that random MAX 2-SAT has a phase structure analogous to the decision problem's. And there is a hope that the maximization problems may help in understanding the decision problems. For 2-SAT this hope is borne out to a degree by our Theorem 6. While our results for $k > 2$ are very limited (see Theorem 15), Conjectures 12 and 14 link the open questions for the maximization and decision thresholds for random satisfiability. At this point we cannot guess the comparative difficulties of resolving the satisfiability threshold conjecture, its maximization analog, or the conjectured link between them.

Our study of random MAX 2-SAT and random MAX CUT was also motivated by recent work on “avoiding a giant component”; we will discuss this in section 8.

We consider several aspects of random MAX 2-SAT and random MAX CUT. We also extend the easiest results to arbitrary CSPs (constraint satisfaction problems).

We will give a second motivation for considering problems of this sort when we take up MAX CUT, in section 8.

2 Notation and model

We write $F(n, m)$ to denote a random 2-SAT formula on n variables, with m clauses. Typically we will fix a constant c and consider $F(n, \lfloor cn \rfloor)$; where it does not matter we will often write cn in lieu of $\lfloor cn \rfloor$ and we often omit the notation $\lfloor \cdot \rfloor$ in other instances. For any formula F , define $\max F$ to be the size of a largest satisfiable subformula of F . Our focus is the functional behavior of $\max F$.

Similarly, we write $G(n, m)$ for a random graph on n vertices with m edges. For any graph G , let \vec{X} describe a partition of the vertices, and let $\text{cut}(G, \vec{X})$ be the number of edges having one vertex in each part of the partition. Define $\max \text{cut}(G) \doteq \max_{\vec{X}} \text{cut}(G, \vec{X})$, and $f_{\text{cut}}(n, m) \doteq \mathbb{E}(\max \text{cut}(G(n, m)))$.

We use standard asymptotic and “order” notation, so for example $f(n) \simeq g(n)$ means $f(n)/g(n) \rightarrow 1$ as $n \rightarrow \infty$, and $f(n) = o(n)$ means $f(n)/n \rightarrow 0$. We will also write $f(n) \lesssim g(n)$ to indicate that f is less than or equal to g *asymptotically* — $\limsup f(n)/g(n) \leq 1$ — though it may be that $f(n) > g(n)$ even for arbitrarily large values of n . Asymptotic results involving two variables, for example concerning 2-SAT formulae on n variables with cn clauses, with c large (or $(1 + \varepsilon)n$ clauses with ε small) should always be interpreted as taking the limit in n second; thus “for any desired error bound there exists a c_0 , such that for all $c > c_0$ there exists an n_0 , such that for all $n > n_0$,” etcetera.

3 Summary of results

We establish several properties of random MAX 2-SAT, random MAX k -SAT, and random MAX CUT, focusing on 2-SAT. This section summarizes our main results and indicates the nature of the proofs; further results and proofs are given in subsequent sections.

One of our goals is to establish the MAX 2-SAT results without depending on those for decision 2-SAT, and in particular to work independently of [BBC⁺01] and reproduce its results. We enjoy some success in this; the

exceptions are our reliance on [BBC⁺01] for the upper bound in Theorem 5 (with an extraneous logarithmic factor arising in the translation), and a more acute form of the same problem in the scaling window, where we lack any corresponding bound for the $\lambda > 1$ case of Theorem 6.

Figure 3 show an “artist’s rendition” of the our results for 2-SAT. For $c < 1$, we expect to satisfy nearly all clauses, while for $c \rightarrow \infty$, we expect to satisfy only about $3/4$ ths of them. The asymptotic behavior for $c < 1$ is understood; so is that for c large (with a log-factor gap in the bounds on the second term); and for $c = 1 \pm \Theta(n^{-1/3})$ (with only a one-sided bound on the second term). We now state these results more exactly; we prove them in the next section.

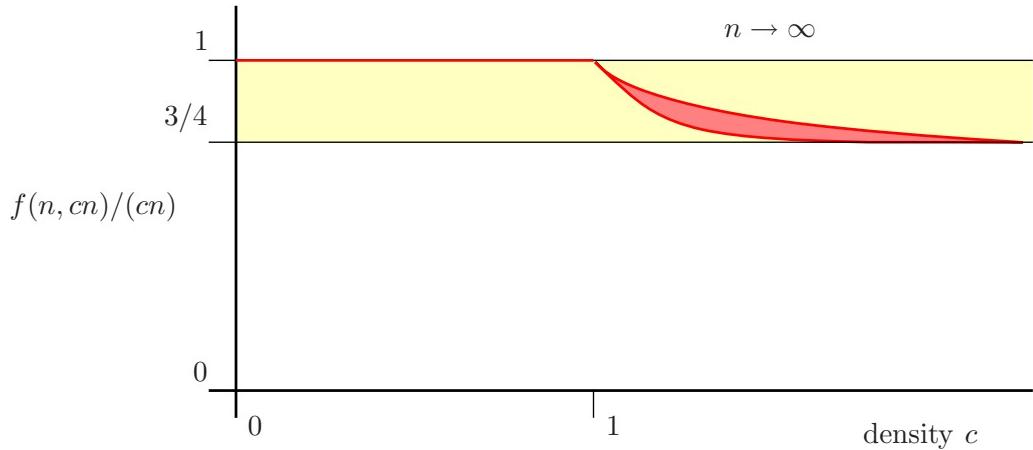


Figure 1: “Artist’s rendition” of the behavior of $f(n, cn)/(cn)$.

For $c < 1$ a random formula $F(n, cn)$ is satisfiable w.h.p., so we would expect $\max F$ to be close to cn in this case; the following theorem shows this to be true.

Theorem 3 *For $c = 1 - \varepsilon$, with any constant $\varepsilon > 0$, $\lfloor cn \rfloor - f(n, \lfloor cn \rfloor) = \Theta(1/(\varepsilon^3 n))$.*

The proof comes from counting the expected number of the “bicycles” shown by [CR92] to be necessary components of an unsatisfiable formula.

For any c , $f(n, cn) \geq \frac{3}{4}cn$, since a random assignment of the variables satisfies each clause with probability $\frac{3}{4}$. The next theorem shows that neither this bound nor the trivial upper bound cn is tight, although for large c , $\frac{3}{4}cn$ is close to correct.

Theorem 4 For c large, $(\sqrt{c} \frac{\sqrt{8}-\sqrt{1}}{3\sqrt{\pi}} - O(1))n \lesssim f(n, cn) - \frac{3}{4}cn \lesssim (\sqrt{c} \sqrt{3 \ln(2)/8})n$.

The values of $\frac{\sqrt{8}-\sqrt{1}}{3\sqrt{\pi}}$ and $\sqrt{3 \ln(2)/8}$ are approximately 0.343859 and 0.509833, respectively. The upper bound is proved by a simple first-moment argument, and the lower bound by analyzing an algorithm; both techniques are exactly those demonstrated in [Spe94, Lecture 6] to analyze the Gale-Berlekamp switching game.

Our next results relate to the low-density case, when c is above but close to the critical value 1. How does $f(n, cn)$ depend on $c = 1 + \varepsilon$ for small ε ?

Theorem 5 For any fixed $\varepsilon > 0$, $(1 + \varepsilon - \varepsilon^3/3)n \lesssim f(n, (1 + \varepsilon)n)$; also, there exist absolute constants α_0 and ε_0 , such that for any fixed $0 < \varepsilon < \varepsilon_0$, $f(n, (1 + \varepsilon)n) \lesssim (1 + \varepsilon - \frac{1}{3}\alpha_0\varepsilon^3/\ln(1/\varepsilon))n$.

That is, a constant fraction of the clauses must remain unsatisfied, but this fraction — $\varepsilon^3/3$ at most — is surprisingly small. The lower bound is proved by using the “differential equation method” (see for example [Wor95]) to exactly analyze a version of the unit-clause heuristic. The upper bound’s proof is a simple first-moment argument; however, for the probability that a sub-formula with density > 1 is satisfiable, it requires the exponentially small bound given by Bollobás et al. [BBC⁺01] (see Theorem 8 below). It is likely that, by replacing our use of [BBC⁺01] with structural properties of the kernel of a sparse random graph, the upper bound’s $\varepsilon^3/\ln(1/\varepsilon)$ can be replaced by ε^3 to match the lower bound up to constants [JS02].

The major significance of [BBC⁺01] was to determine the “scaling window” for random 2-SAT. Without using their result, we prove an analogous result for MAX 2-SAT, and incidentally reproduce most parts of their 2-SAT result.

Theorem 6 Letting $c = 1 + \epsilon = 1 + \lambda n^{-1/3}$, we have

$$\lfloor cn \rfloor - f(n, \lfloor cn \rfloor) = \begin{cases} O(\lambda^3) & \text{if } \lambda > 1; \\ \Theta(1) & \text{if } -1 \leq \lambda \leq 1; \\ \Theta(|\lambda|^{-3}) & \text{if } \lambda < -1. \end{cases}$$

Furthermore, for $\lambda > 1$, for some positive absolute constant, and any $\beta > 0$,

$$\Pr((\lfloor cn \rfloor - f(n, \lfloor cn \rfloor)) > \text{const } \beta \lambda^3) \leq \exp(-\beta \lambda^3).$$

Also,

$$\Pr(F(n, cn) \text{ is satisfiable}) = \begin{cases} \exp(-O(\lambda^3)) & \text{if } \lambda > 1; \\ \Theta(1) & \text{if } -1 \leq \lambda < 1; \\ 1 - \Theta(|\lambda|^{-3}) & \text{if } \lambda < -1. \end{cases}$$

In particular, in the scaling window $c = 1 \pm \lambda n^{-1/3}$, a random formula is satisfiable with probability which is bounded away from 0 and 1 (the exact bounds depending on λ), and it can be made satisfiable by removing a constant-order number of clauses (the constant depending on λ).

In section 6, for MAX k -SAT, we derive analogous results only for c large, reflecting the general state of ignorance regarding the k -SAT phase transition. (For some results on scaling windows for k -SAT see [Wil02].) Still more generally, Theorem 16 describes the high-density case for any MAX CSP. More interestingly, for random MAX k -SAT (including $k = 2$) we observe that $\max F$ is concentrated about its expectation $f(n, cn)$ (as previously remarked in [BFU93]) and that $f(n, cn)/(cn)$ is monotone non-increasing in c . Were $f(n, cn)/(cn)$ also monotone in n , an important property analogous to the satisfiability conjecture would follow; we present this as a conjecture for general MAX CSPs.

In section 7 we consider online versions of MAX 2-SAT, for one of which we prove that a natural greedy algorithm is optimal.

Results for the MAX CUT problem for sparse random graphs, which is closely analogous to random MAX 2-SAT, are presented in section 8.

4 Random MAX 2-SAT

4.1 Sub-critical MAX 2-SAT

One of the most basic facts concerning MAX 2-SAT is that for constants $c < 1$, the expected number of clauses unsatisfied is $o(1)$. This is refined by Theorem 3, which shows the number to be $\Theta(1/(\varepsilon^3 n))$. We now prove the theorem.

THEOREM 3: *Proof.* We write the proof in the SAT equivalent of the “ $G(n, p)$ ” model, because the expressions for the probability of a clause’s presence are cleaner in this model, but adaptation to the $G(n, m)$ model is immediate.

A k -bicycle (see Figure 4.1) is a sequence of clauses $\{\bar{u}, w_1\}$, $\{\bar{w}_1, w_2\}$, \dots , $\{\bar{w}_k, v\}$ where literals w_1, w_2, \dots, w_k are distinct as variables (none is

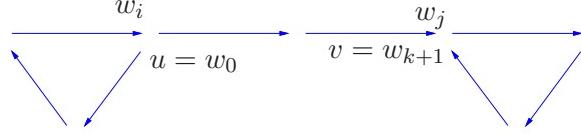


Figure 2: Sequence of clause-derived implications for a bicycle. Start the walk from u , proceed clockwise to w_i (which equals either u or \bar{u}), continue right to w_j , and again go clockwise to terminate at v (which equals either w_j or \bar{w}_j).

the same as nor the complement of another) and $u \in \{w_i, \bar{w}_i\}$, $v \in \{w_j, \bar{w}_j\}$ for some $1 \leq i, j \leq k$. (Think of it as a “walk” in which the first and last variables are also both visited *en route*.) Because satisfying a clause $\{\bar{u}, v\}$ means that if u is true then v must be true, such a clause yields an implication $u \rightarrow v$ (and a complementary implication $\bar{v} \rightarrow \bar{u}$); Figure 4.1 represents such a sequence of implications for a bicycle. Chvátal and Reed [CR92] argue that if a formula is infeasible then it contains a bicycle. Thus if we delete an edge from every bicycle, the remaining subformula is satisfiable.

The number of potential k -bicycles, whether or not present in a given formula F , is at most $k^2(2n)^k$. The probability that all $k+1$ clauses of a given bicycle are present in a random formula F is at most $[(cn)/(2^2 \binom{n}{2})]^{k+1} = [c/(2(n-1))]^{k+1}$, so the expected number of k -bicycles is $\lesssim k^2 c^{k+1}/(2n)$. If we delete one edge in every bicycle, we obtain a satisfiable formula. For any fixed $c < 1$, $\sum_{k=1}^n k^2 c^{k+1}/(2n) = \Theta(1/n)$. Thus, the expected number of edges we need to delete is at most $O(1/n)$ and $f(n, \lfloor cn \rfloor) \geq \lfloor cn \rfloor - O(1/n)$.

To obtain the lower bound we show that with probability at least $\Theta(1/(\varepsilon^3 n))$ the formula F is not satisfiable. This clearly implies an upper bound $f(n, \lfloor cn \rfloor) \leq \lfloor cn \rfloor - \Theta(1/(\varepsilon^3 n))$. To this goal we employ the second moment method.

For simplicity here, we will restrict ourselves to 3-bicycles, which will only establish “ $\Theta_\varepsilon(1/n)$ ”, that is, something of order $\Theta(1/n)$ but with hidden constants that may depend on ε . The full proof is the same but using bicycles of lengths up to $1/\varepsilon$, not just length 3, and parallels the proof of Theorem 6, case $\lambda \leq -1$. (In fact, taking $\lambda = \lambda(n) = \varepsilon n^{1/3}$ there establishes the current theorem completely.)

Consider 4-tuples of clauses of the form $\{\bar{u}_1, u_2\}, \{\bar{u}_2, \bar{u}_1\}, \{u_1, u_3\}, \{\bar{u}_3, u_1\}$, where u_1, u_2, u_3 are arbitrary variables. One observes that this sequence of clauses is a 3-bicycle, and, moreover, its presence in the random formula F implies non-satisfiability. We now show, using second moment

method, that the number B_3 of such bicycles is at least one with probability at least $\Omega(1/n)$. We have $\mathbb{E}[B_3^2] = \sum \mathbb{P}(X \in F, X' \in F)$, where the sum runs of over the pairs of 3-bicycles X, X' of the form above, and $X \in F$ means all the clauses of X are present in F . We decompose the sum into three parts: the sum over pairs X, X' with $X = X'$, the sum over pairs that do not have common clauses and the rest. It is easy to see that the first sum is simply $\mathbb{E}[B_3]$ which is $\Theta(1/n)$, by the argument for upper bound. To analyze the second sum note that for each fixed pair X, X' with no common clauses, we have $\mathbb{P}(X, X' \in F) = \mathbb{P}(X \in F)\mathbb{P}(X' \in F)$, when replacement of clauses is allowed. (When replacement is not allowed the reader can check that the difference between the left and the right-hand sides is very small, and the rest of the argument goes through). Then, this sum is smaller than $\sum_{X, X'} \mathbb{P}(X \in F)\mathbb{P}(X' \in F) = (\mathbb{E}[B_3])^2$, where the sum now runs over all the pairs X, X' . For the third sum we have two cases. First case is pairs X, X' defined on the same set of variables. For example $X = \{\bar{u}_1, u_2\}, \{\bar{u}_2, \bar{u}_1\}, \{u_1, u_3\}, \{\bar{u}_3, u_1\}$ and $X' = \{\bar{u}_1, u_2\}, \{u_2, u_1\}, \{\bar{u}_2, u_3\}, \{\bar{u}_3, \bar{u}_2\}$, share one clause $\{\bar{u}_1, u_2\}$ and are defined over the same set of variables. There are $O(n^3)$ choices for the variables u_1, u_2, u_3 in these pairs. But since $X \neq X'$ then there are altogether at least five clauses in X and X' together. For a given pair, the probability that all these clauses are present in F is $O(1/n^5)$. Then the expected number of such pairs $X, X' \in F$ is $O(1/n^2) = o(1/n)$.

The second case is pairs X, X' defined over different set of variables. Since they share a clause then the pair is defined on exactly four variables. But then there are at least six clauses in this pair. We obtain that the expected number of such pairs X, X' which belong to F is at most $O(n^4)O(1/n^6) = O(1/n^2) = o(1/n)$.

We conclude that $\mathbb{E}[B_3^2] = \mathbb{E}[B_3] + o(1/n) = \Theta(1/n) + o(1/n)$. We now use the bound $\mathbb{P}(Z \geq 1) \geq (\mathbb{E}[Z])^2/\mathbb{E}[Z^2]$, which holds for any non-negative integer random variable Z . Applying this bound to B_3 we obtain $\mathbb{P}(B_3 \geq 1) \geq (\mathbb{E}[B_3])^2/\mathbb{E}[B_3^2] \geq \Theta(1/n^2)/(\Theta(1/n) + o(1/n)) = \Theta(1/n)$. This completes the proof. \square

It is worth pointing out the following simple fact, upon which we will shortly improve.

Remark 7 For $c > 1$, $f(n, cn) \gtrsim n(\frac{3}{4}c + \frac{1}{4})$.

Proof. It suffices to show that for any $\varepsilon > 0$, for all n sufficiently large, $f(n, cn) \geq (\frac{3}{4}c + \frac{1}{4} - \varepsilon)n$. Select the first $(1 - \varepsilon)n$ clauses, and let \vec{X} be a best assignment for it. By Theorem 3, \vec{X} satisfies an expected $(1 - \varepsilon)n - o(1)$ of

these first clauses. Also, an expected $3/4$ ths of the remaining $(c - 1 + \varepsilon)n$ clauses are satisfied, yielding the claim. \square

4.2 High-density random MAX 2-SAT

While it is well known that for $c > 1$, $F(n, cn)$ is a.s. unsatisfiable, is it possible that even for c large, *almost* all clauses are satisfiable? Theorem 4 rules this out by showing that a constant fraction of clauses must go unsatisfied; up to a constant, it also provides a matching lower bound.

THEOREM 4: *Proof of the upper bound.* The proof is by the first-moment method. If $\max F > (1 - r)cn$ then there is a satisfying assignment of a subformula F' which omits rcn or fewer clauses, and where (taking F' to be maximal) all the omitted clauses are unsatisfied. Any fixed assignment satisfies each (random) clause of F' w.p. $3/4$ and unsatisfies each omitted clause w.p. $1/4$, so by linearity of expectations, the probability that there exists such an F' is

$$P = \mathbb{P}(\exists \text{ satisfiable } F') \leq 2^n \sum_{k=0}^{rcn} \binom{cn}{k} \left(\frac{3}{4}\right)^{cn-k} \left(\frac{1}{4}\right)^k.$$

For $r < \frac{1}{4}$ the sum is dominated by the last term. From Stirling's formula $n! \simeq \sqrt{2\pi n} (n/e)^n$,

$$(1) \quad \binom{cn}{rcn} \simeq 1/\sqrt{2\pi r(1-r)cn} \quad (r^{-r}(1-r)^{-(1-r)})^{cn}.$$

Substituting (1) into the previous expression,

$$P \lesssim 1/\sqrt{2\pi r(1-r)cn} 2^n \quad (r^{-r}(1-r)^{-(1-r)}(3/4)^{1-r}(1/4)^r)^{cn}.$$

Substituting $r = 1/4 - \varepsilon$,

$$\frac{1}{cn} \ln P \lesssim \ln(2)/c - (8/3)\varepsilon^2 + O(\varepsilon^3),$$

so that for $\varepsilon > \sqrt{(3/8) \ln 2/c}$, as $n \rightarrow \infty$, $P \rightarrow 0$. The conclusion follows.

Proof of the lower bound. The proof is algorithmic. When variables X_1, \dots, X_k have been set, define the reduced formula F_k in which any clause

containing a True literal is removed and “scored”, and False literals are removed from the remaining clauses. (Clauses with 0 variables remaining are permanently unsatisfied.) Define a potential function $q(F_k)$ to be the number of clauses already satisfied, plus $3/4$ the number of 2-variable clauses (“2-clauses”), plus $1/2$ the number of 1-variable clauses (“unit clauses”). Note that randomly assigning the remaining variables satisfies an expected total number of clauses precisely $q(F_k)$, so q is a lower bound on the number of clauses satisfiable.

After variables X_1, \dots, X_{k-1} have been set to define F_{k-1} , our algorithm sets X_k in whichever of the two ways gives an F_k with larger value $q(F_k)$. (Ties may be broken arbitrarily.) In F_{k-1} , let the number of appearances of X_k and \bar{X}_k in unit clauses be denoted by A_1 and \bar{A}_1 , and their number of appearances in 2-clauses by A_2 and \bar{A}_2 . If X_k is set to True, then

$$q(F_k) - q(F_{k-1}) = \Delta_k \doteq \frac{1}{2}(A_1 - \bar{A}_1) + \frac{1}{4}(A_2 - \bar{A}_2),$$

and if X_k is set False, then $q(F_k) - q(F_{k-1}) = -\Delta_k$. Note that $q(F_k) = q(F_{k-1}) + |\Delta_k|$ is a lower bound on the number of satisfiable clauses, and $q(F_0) = \frac{3}{4}cn$.

With $k-1$ variables already set, F_{k-1} a.s. has $(\frac{1}{2} \pm O(1/\sqrt{c}))2\frac{k-1}{n}\frac{n-k+1}{n}$. cn unit clauses, and $(\frac{n-k+1}{n})^2 \cdot cn$ 2-clauses, on the remaining variables. (The reason for $\frac{1}{2} \pm O(1/\sqrt{c})$ instead of $\frac{1}{2}$ is that we set the previous variables in a biased manner.) Also, conditioned on the number of clauses, F_{k-1} is a uniformly random formula (each “slot” being equally likely to be filled by any of the remaining literals). For n large, A_1 and \bar{A}_1 are approximated by independent Poisson random variables with parameter $(\frac{1}{2} \pm O(1/\sqrt{c}))\frac{k-1}{n}c$, and A_2 and \bar{A}_2 by Poissons with parameter $\frac{n-k+1}{n}c$. By assumption, c is large, so each of these distributions is approximately Gaussian, and their sum Δ_k is also approximately Gaussian, with mean 0 (by symmetry) and variance

$$\begin{aligned} \sigma_k^2 &= 2 \cdot (\frac{1}{2})^2 \cdot \text{Var}(A_1) + 2 \cdot (\frac{1}{4})^2 \cdot \text{Var}(A_2) \\ &= c \left((\frac{1}{4} \pm O(1/\sqrt{c})) \frac{k-1}{n} + \frac{1}{8} \frac{n-k+1}{n} \right). \end{aligned}$$

For $Z \sim N(0, 1)$, it is well known that $\mathbb{E}|Z| = \sqrt{2/\pi}$; thus $\mathbb{E}|\Delta_k| =$

$\sqrt{2/\pi} \sigma_k = \sqrt{2/\pi} \sqrt{c(\frac{1}{4}\frac{k-1}{n} + \frac{1}{8}\frac{n-k+1}{n})} \pm O(1)$. Finally,

$$\begin{aligned} \mathbb{E}(q(F_n)) &\geq \frac{3}{4}cn + \sum_{k=0}^{n-1} \mathbb{E}(|\Delta_k|) \\ &\approx \frac{3}{4}cn + \int_0^n \mathbb{E}(|\Delta_k|) dk \\ &\gtrsim \frac{3}{4}cn + \left(\sqrt{c} \frac{\sqrt{8} - \sqrt{1}}{3\sqrt{\pi}} - O(1) \right) n. \end{aligned}$$

□

We remark that in the preceding proof, X_k was set True or False so as to maximize half the number of satisfied unit clauses plus a quarter the number of satisfied 2-clauses. This is reminiscent of the “policies” in [AS00]. There, the goal was to satisfy as a dense a 3-SAT formula as possible; unit clauses always had to be satisfied, and variables were set so as to maximize a linear combination of the number of satisfied 2-clauses and 3-clauses. In [AS00], the linear combination which was optimal for the purpose changed during the course of the algorithm; the determination of the optimal combinations, and the proof of optimality, was a main result of the paper. In the present case, though, it is evident that the ratio 1:2 is optimal: for c large, the potential function q predicts the expected number of clauses satisfiable almost exactly. The difference can be ascribed to the fact that here c is “large”, and in [AS00] the corresponding parameter (the initial 3-clause density) was fixed (relevant values were in the range of 3.145 to 3.26). Were we to try to tune the MAX 2-SAT algorithm above for small values of c , more complex methods like those of [AS00] would presumably be needed.

4.3 Low-density random MAX 2-SAT

For low-density formulas, with $c = 1 + \varepsilon$ and $\varepsilon > 0$ a small constant, the bounds of Theorem 4 are inapplicable. It is still true (from Remark 7) that we expect to satisfy at least $(1 + \frac{3}{4}\varepsilon)n$ clauses, but it is not obvious whether the best answer is this, or close to the full number of clauses $(1 + \varepsilon)n$, or something in between. In this section we prove Theorem 5 which shows that $(1 + \varepsilon)n - f(n, cn)$, the number of clauses we must dissatisfy, lies between $\Theta(\varepsilon^3 n / \ln(1/\varepsilon))$ and $\Theta(\varepsilon^3 n)$. That is, a linear fraction of clauses must be rejected, but this fraction, at most $\Theta(\varepsilon^3)$, is surprisingly small. We will employ the following theorem of Bollobás et al. [BBC⁺01] on random 2-SAT.

Theorem 8 ([BBC⁺01], Corollary 1.5) *There exist positive constants α_0 and ε_0 such that for any $0 < \varepsilon < \varepsilon_0$ and sufficiently large n , $\mathbb{P}[F(n, (1 + \varepsilon)n)] \leq \exp(-\alpha_0\varepsilon^3n)$.*

(Here, α_0 is the \liminf of the constant implicit in Θ in the theorem in [BBC⁺01].) The $\exp(-\Theta(\varepsilon^3n))$ probability of satisfiability in random 2-SAT translates into an expected $O(\varepsilon^3n/\ln(1/\varepsilon))$ unsatisfied clauses in random MAX 2-SAT.

THEOREM 5: *Proof of the upper bound.* The proof is by the first-moment method. Let $c = 1 + \varepsilon$. Let F' range over subformulas of F which omit rcn or fewer clauses. Specifying $r < 1/4$, the conditions of Theorem 8 apply, so

$$(2) \quad P = \mathbb{P}(\exists \text{ maximally satisfiable } F') \leq \sum_{k=0}^{rcn} \binom{cn}{k} \left(\frac{1}{4}\right)^{rcn} e^{-\alpha_0(\varepsilon - \frac{k}{n})^3 n};$$

as $r < 1/4$, the sum is dominated by the last term. Using (1) to approximate $\binom{cn}{crn}$,

$$\frac{1}{cn} \ln P \lesssim -r \ln r - (1-r) \ln(1-r) - \alpha_0(\varepsilon - cr)^3/c - r \ln(4).$$

First observe that as $\varepsilon \rightarrow 0$, for any $r = o(\varepsilon)$, this is

$$= -r \ln r(1 + o(1)) - \alpha_0 \varepsilon^3(1 + o(1)) - r \ln(4).$$

For any constant $b < 1/3$, if $r = b\alpha_0\varepsilon^3/\ln(1/\varepsilon)$, this is

$$= 3b\alpha_0\varepsilon^3(1 + o(1)) - \alpha_0\varepsilon^3(1 + o(1)) < 0.$$

That is, it is unlikely that asymptotically fewer than $(1/3)\alpha_0\varepsilon^3/\ln(1/\varepsilon)$ clauses can go unsatisfied.

Proof of the lower bound. The proof is algorithmic, and of the sort familiar from [AS00] and previous works. It analyzes a version of the “unit-clause” heuristic. Initially, “seed” the algorithm by randomly deleting a variable from each of, say, $n^{1/10}$ random 2-clauses to convert them to unit clauses. While F has any unit clauses, select one at random and set its variable to satisfy the clause. Continue until no unit clauses remain. The analysis consists of counting the clauses unsatisfied in these steps, and justifying the

assertion that when there are no more unit clauses, $o(1)$ further clauses need be unsatisfied.

When k variables have been set, let the number of 2-clauses be denoted $m_2(k)$, the number of unit clauses $m_1(k)$, and the number of unset variables $m(k) = n - k$. In one step, the changes in these quantities are $\Delta m = -1$, $\mathbb{E}(\Delta m_2) = -\frac{2}{m}m_2$, and $\mathbb{E}(\Delta m_1) = -1 - \frac{1}{m}m_1 + \frac{1}{m}m_2$ (assuming that $m_1 > 0$ before the step). Over a large number of steps, the net changes will be a.s. a.e. equal to the expectations. Renormalizing with $\rho = m/n$, $\rho_1 = m_1/n$, and $\rho_2 = m_2/n$, the differential equation method (see for example [AS00, Wor95]) asserts that (ρ_1, ρ_2) a.s. a.e. obey the differential equations

$$d\rho_2/d\rho = \frac{2\rho_2}{\rho} \quad d\rho_1/d\rho = 1 + \frac{\rho_1}{\rho} - \frac{\rho_2}{\rho}.$$

With boundary conditions that for $\rho = 1$ (i.e., initially), $\rho_2 = c$ and $\rho_1 = 0$, the unique solution is

$$\rho_2 = c\rho^2 \quad \rho_1 = c\rho - c\rho^2 + \rho \ln \rho.$$

This results in $\rho_1 = 0$ at two times: initially, when $\rho = 1$, and also for $\rho = \rho^*$ satisfying

$$(3) \quad c = \ln(\rho^*)/(\rho^* - 1).$$

While $\rho > \rho^*$, the only clauses ever unsatisfied are unit clauses which contain the negation of the variable being set, and the expected number of such rejected clauses per step is $\frac{1}{2m}m_1 = \frac{\rho_1}{2\rho}$. Integrating over the period ρ^* to 1,

$$\begin{aligned} \int_{\rho^*}^1 \frac{\rho_1}{2\rho} d\rho &= \frac{1}{2} \int_{\rho^*}^1 (c - c\rho + \ln \rho) d\rho \\ &= \frac{1}{2} \left(c\rho - c\rho^2/2 + \rho \ln \rho - \rho \right) \Big|_{\rho^*}^1 \end{aligned}$$

which, substituting for c from (3)

$$(4) \quad = \frac{1}{2}(\rho^* - 1) - \frac{1}{4}(\rho^* + 1) \ln \rho^*.$$

So from $\rho = 1$ to $\rho = \rho^*$, the number of clauses dissatisfied by the algorithm is a.s. a.e. n times expression (4). After this time, the remaining (uniformly random) 2-SAT formula has density $\rho_2(\rho^*) / \rho^* = c\rho^* =$

$\ln(\rho^*)\rho^*/(\rho^* - 1) < 1$ since $\ln(\rho^*) < \rho^* - 1$ and $\rho^* < 1$, and thus (by Theorem 7) contributes $o(1)$ to the expected number of unsatisfied clauses. In short, the algorithm a.s. fails to satisfy a.e. $(\frac{1}{2}(\rho^* - 1) - \frac{1}{4}(\rho^* + 1)\ln\rho^*)n$ clauses. For ρ^* (asymptotically) close to 1, the number of dissatisfied clauses is $\simeq n(1 - \rho^*)^3/24$. In particular, with $\varepsilon > 0$ asymptotically small and $c = 1 + \varepsilon$, $\rho^* \simeq 1 - 2\varepsilon$, and the number of dissatisfied clauses is $\simeq n\varepsilon^3/3$.

□

Two remarks. First, in addition to the asymptote, the proof gives a precise parametric relationship (as functions of ρ^*) between the clause density c (given by (3)) and the rejected-clause density (given by (4)). Solving numerically, for $c = 1.5$ we find rejected-clause density ≈ 0.0183275 , and for $c = 2$ — where naively the rejected-clause density would be $\frac{1}{4}c = 0.5$ — we achieve rejected-clause density ≈ 0.0809517 .

Second, with the solution in hand, the asymptotic behavior is easy to see without the need for differential equations. This alternate proof is not fully rigorous, but is more intuitive and more robust; it is the basis of the analysis within the scaling window (see Theorem 6).

THEOREM 5: *Alternate proof of lower bound.* Consider what happens when $m = (1 - \delta)n$ variables remain unset. The number of 2-clauses is a.s. $m_2 \simeq (1 - \delta)^2(1 + \varepsilon)n \simeq (1 + \varepsilon - 2\delta)n$. The expected increase in the number of unit clauses is then $\mathbb{E}(\Delta m_1) = -1 - m_1/m + m_2/m \geq -1 + m_2/m$ (and the neglected m_1/m is not only conservative, but will also prove to be insignificantly small). Thus, $\mathbb{E}(\Delta m_1) \geq -1 + [(1 + \varepsilon - 2\delta)n]/[(1 - \delta)n] \simeq \varepsilon - \delta$. At $\delta = 0$, the number of unit clauses increases by ε per step, this increase linearly falls to 0 per step by $\delta = \varepsilon$, and further to $-\varepsilon$ by $\delta = 2\varepsilon$: the expected number of unit clauses is bounded by an inverted parabola, with base $2\varepsilon n$ and height $\frac{1}{2}\varepsilon^2 n$. At each step about $1/(2n)$ th of the unit clauses get dissatisfied. The area under the parabola, times this $1/(2n)$ factor, is $\frac{2}{3} \cdot \text{base} \cdot \text{height} \cdot 1/(2n) = \frac{1}{3}\varepsilon^3 n$.

□

5 The MAX 2-SAT scaling window

For random MAX 2-SAT, we have seen that for fixed $c < 1$, $\lfloor cn \rfloor - f(n, \lfloor cn \rfloor) = \Theta(1/n)$, and for $c > 1$, $cn - f(n, cn) = \Theta(n)$. That is, random MAX 2-SAT experiences a phase transition around $c = 1$. It is natural to ask about the scaling window around the critical threshold: What is the interval around $c = 1$ within which $\lfloor cn \rfloor - f(n, \lfloor cn \rfloor) = \Theta(1)$? Theorem 6 shows that the scaling window is $c = 1 \pm \Theta(n^{-1/3})$. The corresponding question

for random 2-SAT is the range in which $\mathbb{P}(F(n, \lfloor cn \rfloor))$ is satisfiable) = $\Theta(1)$. This was shown by [BBC⁺01] to be $c = 1 \pm \Theta(n^{-1/3})$ with their result reproduced as Theorem 9 here.

Theorem 9 (Bollobás et al, [BBC⁺01]) *Let $F(n, cn)$ be a random 2-SAT formula, with $c = 1 + \lambda_n n^{-1/3}$. There are absolute constants $0 < \varepsilon_0 < 1$, $0 < \lambda_0 < \infty$, such that the probability F is satisfiable is: $1 - \Theta(1/|\lambda_n|^3)$, when $-\varepsilon_0 n^{1/3} \leq \lambda_n \leq -\lambda_0$; $\Theta(1)$, when $-\lambda_0 \leq \lambda_n \leq \lambda_0$; and $e^{-\Theta(\lambda_n^3)}$, when $\lambda_0 \leq \lambda_n \leq \varepsilon_0 n^{1/3}$.*

That the two scaling windows are the same is no coincidence, and in fact Theorem 6 reestablishes much of Theorem 9 independently.

THEOREM 6: *Proof.* Note that, provided we prove the bounds for the cases $\lambda \leq -1$ and $\lambda \geq 1$, the bound for the case $|\lambda| < 1$ follows immediately, since we obtain that the probability of satisfiability is at least $\exp(-O(\lambda^3)) \geq \exp(-O(1))$ and at most $1 - \Theta(1/|\lambda|^3) \leq 1 - \Theta(1)$, where in both cases $|\lambda| < 1$ was used. The more interesting cases $|\lambda| \geq 1$ are considered in two subsections below.

5.1 Case $c = 1 + \lambda n^{-1/3}$, $\lambda \leq -1$

For convenience we write $c = 1 - \lambda n^{-1/3}$ and $\lambda \geq 1$. The proof for this case is very similar to that of Theorem 3 and uses the notion of bicycles. (As in the earlier case, we work in the equivalent of the $G(n, p)$ model for notational convenience, with the understanding that the proof works equally well in the $G(n, m)$ model.) As before, the number of clauses that must be dissatisfied is bounded by the number of bicycles. The expected number of k -bicycles is at most $k^2 c^{k+1}/(2n) = k^2 (1 - \lambda n^{-1/3})^{k+1}/(2n)$. Using the formula $\sum_{k \geq 1} k^2 \rho^k = \frac{\rho(1+\rho)}{(1-\rho)^3}$ which for $\rho \simeq 1$ is $\simeq 2/(1-\rho)^3$, we have

$$(5) \quad \sum_{1 \leq k < \infty} k^2 (1 - \lambda n^{-1/3})^{k+1}/(2n) \simeq 2/\lambda^3.$$

Therefore $\lfloor cn \rfloor - f(n, \lfloor cn \rfloor) = O(1/\lambda^3)$. Using Markov's inequality we also obtain that the probability that the formula is unsatisfiable is at most the expected number of bicycles, that is, at most $O(1/\lambda^3)$.

We now obtain a matching lower bound. Consider only “bad” bicycles, in which $u = \bar{w}_i$, $v = \bar{w}_j$, and $i < j$. Note that no bad bicycle is completely satisfiable, since the first “wheel” $u \rightarrow \dots \rightarrow w_i = \bar{u}$ requires $u = \text{False}$ and thus $w_i = \text{True}$; whereupon the path (technically called the “top tube”

of a bicycle) $w_i \rightarrow \dots \rightarrow w_j$ implies $w_j = \text{True}$; and the second wheel $w_j \rightarrow \dots \rightarrow v = \bar{w}_j$ provides a contradiction. Note that about 1/8th of the potential bicycles are bad.

Let B_k denote the number of the bad k -bicycles. Since

$$(6) \quad \begin{aligned} \mathbb{E}(\#\text{unsatisfiable clauses}) &\geq \Pr(F \text{ unsatisfiable}) \\ &\geq \Pr\left(\sum_{k \leq K} B_k \geq 1\right), \end{aligned}$$

it suffices to prove that this is

$$(7) \quad = \Omega(1/\lambda^3);$$

we will show this for $K = (1/\lambda)n^{1/3}$. Repeating the argument for (5), we obtain that

$$\mathbb{E}\left[\sum_{k \leq K} B_k\right] \gtrsim (2/(8e))/\lambda^3,$$

the $1/(8e)$ coming from the series' truncation at K and the use of only bad bicycles. To obtain (7) it suffices prove that

$$(8) \quad \mathbb{E}\left[\left(\sum_{k \leq K} B_k\right)^2\right] = (1 + O(1)) \cdot \mathbb{E}\left[\sum_{k \leq K} B_k\right],$$

for then

$$\mathbb{P}\left(\sum_{k \leq K} B_k \geq 1\right) \geq \frac{\mathbb{E}\left[\sum_{k \leq K} B_k\right]^2}{\mathbb{E}\left[\left(\sum_{k \leq K} B_k\right)^2\right]} = \frac{\mathbb{E}\left[\sum_{k \leq K} B_k\right]^2}{\mathbb{E}\left[\sum_{k \leq K} B_k\right](1 + O(1))} = \Omega(1/\lambda^3).$$

We will prove (8) with $O(1/\lambda^3)$ filling in for $O(1)$ (recall that $\lambda \geq 1$). Consider pairs of k, k' -bicycles X, X' with $k, k' \leq K$. It suffices to show that for every X ,

$$(9) \quad \sum_{X' \neq X} \mathbb{P}(X' \subseteq F | X \subseteq F) = O(1/\lambda^3),$$

because then

$$\begin{aligned} \mathbb{E}\left[\left(\sum_k B_k\right)^2\right] &= \sum_{X, X'} \mathbb{P}(X, X' \subseteq F) \\ &= \sum_X \Pr(X \subseteq F) [1 + \sum_{X' \neq X} \Pr(X' \subseteq F | X \subseteq F)] \\ &\leq \mathbb{E}\left[\sum_k B_k\right](1 + O(1/\lambda^3)). \end{aligned}$$

Establishing (9) is the nub of the proof. First, observe that for any bicycle X' sharing no literals with X , $\Pr(X' \subseteq F \mid X \subseteq F) \leq \Pr(X' \subseteq F)$, and so such bicycles X' contribute $\leq \mathbb{E} \sum_k B_k = O(1/\lambda^3)$ to the sum.

Given a bicycle $X' = \{u, w_1\}, \{\bar{w}_1, w_2\}, \dots, \{\bar{w}_k, v\}$, a sequence of literals w_i, w_{i+1}, \dots, w_j from X' is defined to be a type I excursion if literals w_i, w_j belong to X but literals w_{i+1}, \dots, w_{j-1} do not. (If $j = i + 1$, a sequence w_i, w_{i+1} is a type I excursion if the corresponding clause $(\bar{w}_i, w_{i+1}) \in X'$ does not belong to X .) A sequence of literals u', w_1, \dots, w_j in X' is defined to be a type II excursion if the literal w_j belongs to X , but u, w_1, \dots, w_{j-1} do not. Similarly, a sequence w_j, w_{j+1}, \dots, v' in X' is defined to be a type III excursion.

Bicycles X' which are neither equal to X nor disjoint from X must have at least one excursion (and at most one each of excursions of type II and III). It suffices to establish (9) for such bicycles X' . We will just show that the expected number of bicycles X' with one type II excursion, no type III excursion, and any number $r \geq 0$ of type I excursions, is $O(1/\lambda^3)$; the other three cases (classified by the number of type II and III excursions) follow similarly.

Since a collection of excursions uniquely defines X' , it is enough to count such collections. Let the lengths of the type I excursions be $m_1, m_2, \dots, m_r \geq 2$ and that of the type II excursion m_{II} , where the length is defined by the number of literals.

For each type I excursion there are two endpoints (literals) which belong to X . Since the size of X is $\leq K = (1/\lambda)n^{1/3}$, there are $\leq K^{2r} = (1/\lambda^{2r})n^{2r/3}$ choices for all the end points. The i th type I excursion contains $m_i - 2$ literals not from X , so there are at most $(2n)^{m_i-2}$ ways of selecting them. The excursion contains $m_i - 1$ clauses, all not from X , so the probability they are all present in F is $(1 - \lambda n^{-1/3})^{m_i-1}/(2n)^{m_i-1}$.

Similarly, for the type II excursion, there are at most K choices for the endpoint literal w_{j-1} , which belongs to X , and at most $(2n)^{m_{II}-2}$ choices for other literals u', w_1, \dots, w_{j-2} . The excursion contains $m_{II} - 1$ clauses, all not from X , so the probability that they are all present in F is $(1 - \lambda n^{-1/3})^{m_{II}-1}/(2n)^{m_{II}-1}$.

Combining, we obtain that the expected number of bicycles X' containing exactly r type I excursions, one type II excursion, and no type III

excursions is

$$\begin{aligned} \#(r, 0, 1) &\leqslant \sum_{m_{II}, m_1, \dots, m_r \geqslant 2} (1/\lambda^{2r+2}) n^{\frac{2r+2}{3}} (2n)^{m_{II}-2+\sum_i m_i - 2r} \\ &\quad \times \frac{(1-\lambda n^{-1/3})^{m_{II}-1+\sum_i m_i - r}}{(2n)^{m_{II}-1+\sum_i m_i - r}} \\ &= \frac{1}{2^{r+1} \lambda^{2r+2} n^{\frac{r+1}{3}}} \sum_{m_{II}, m_1, \dots, m_r \geqslant 2} (1-\lambda n^{-1/3})^{m_{II}+\sum_i m_i - r - 1}. \end{aligned}$$

Note that

$$\begin{aligned} &\sum_{m_{II}, m_1, \dots, m_r \geqslant 2} (1-\lambda n^{-1/3})^{m_{II}+\sum_i m_i - r - 1} \\ &= \sum_{m_{II}, m_1, \dots, m_r \geqslant 1} (1-\lambda n^{-1/3})^{m_{II}+\sum_i m_i} \\ &= \left(\sum_{m \geqslant 1} (1-\lambda n^{-1/3})^m \right)^{r+1} \\ &\leqslant \frac{1}{\lambda^{r+1} n^{-\frac{r+1}{3}}}. \end{aligned}$$

Applying this to the equality above we obtain

$$\begin{aligned} \#(r, 0, 1) &\leqslant \frac{1}{2^{r+1} \lambda^{3r+3}}, \text{ and} \\ \sum_{r \geqslant 0} \#(r, 0, 1) &\leqslant \frac{1}{2\lambda^3(1-2\lambda^3)} = O(1/\lambda^3). \end{aligned}$$

With similar calculations for $\#(r, \cdot, \cdot)$ this establishes (9), and completes the proof of the case $\lambda \leqslant -1$ of Theorem 6. \square

5.2 Case $c = 1 + \lambda n^{-1/3}$, $\lambda \geqslant 1$

The proof of this part resembles the alternate proof of Theorem 5. There we showed that $m_1(t)$ a.s. a.e. followed a parabolic trajectory. Both there and here, at time $t = \varepsilon n$, the expectation given by the parabola is $\frac{1}{2}\varepsilon^2 n$, and the typical deviations (the standard deviation) from summing εn binomial r.v.s with distributions near to $B(n, 1/n)$ is about $\sqrt{\varepsilon n}$.

In the previous case, with $\varepsilon = \Theta(1)$, the deviations were a.s. tiny compared with the expectation, but here, with $\varepsilon = \lambda n^{-1/3}$, the standard deviation of $\sqrt{\lambda} n^{1/3}$ is of the same order (in terms of n) as the expectation

of $\frac{1}{2}\lambda^2 n^{1/3}$: the trajectory is not predictable in an a.s. a.e. sense. Figure 3 shows two typical samples (with $\lambda = 2$ and $n = 10,000$) against the nominal parabolic trajectory. The analysis is thus more involved.

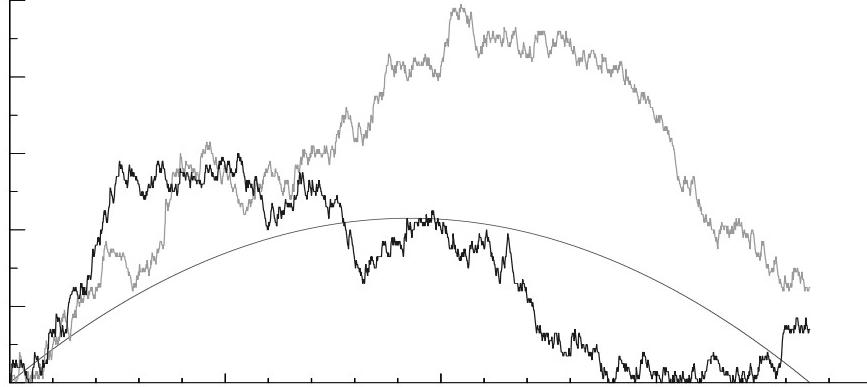


Figure 3: Nominal parabolic trajectory of $m_1(t)$ vs t , and two random samples for density $1 + \lambda n^{-1/3}$ ($\lambda = 2$, $n = 10,000$). With density $1 + \Theta(n^{-1/3})$, the random fluctuations are of the same order as the nominal values.

As before, we analyze the unit-clause resolution algorithm in which if there are any unit clauses (if $m_1(t) > 0$) we choose one at random and set its literal True, and otherwise we choose a random literal (from the variables not already set) and set it True.

Our analysis proceeds in three phases. Phase I proceeds until time $T = 2\varepsilon n$, and we show that in this period, there is an exponentially small chance that m_1 is ever much larger than its expectation. In Phase II, we continue unit-clause resolution until $m_1(t) = 0$; we show that this happens quickly. These will give the required bounds on the integrated number of unit clauses, and in turn unsatisfied clauses, produced by unit-clause resolution. In Phase III we have a formula of density $\leq 1 - \varepsilon n$, and we simply apply the (non-algorithmic) proof of the Theorem's case $\lambda \leq -1$, proved in Section 5.2.

5.2.1 Useful facts

We first establish a simple relation, useful for Phase I and essential for Phase II. The number of 2-clauses remaining (both of whose variables remain) at time δn is $m_2(\delta n) \sim B(n(1 + \varepsilon), (1 - \delta)^2)$. Thus for all times

$t \leq \frac{1}{2}n(1 + \varepsilon)$ (much longer than the times $\Theta(\varepsilon n)$ in which we are interested),

$$(10) \quad \Pr \left(\max_{\delta \leq \frac{1}{2}} \left[m_2(\delta n) - n(1 + \varepsilon)(1 - \delta)^2 \geq n^{3/5} \right] \right) \leq \exp(-\Theta(n^{1/5})).$$

We prove (10) using the Chernoff bound that for a sum X of independent 0-1 Bernoulli random variables with parameters p_1, \dots, p_n and expectation $\mu = \sum_{i=1}^n p_i$,

$$(11) \quad \Pr(X \geq \mu + \Delta) \leq \exp(-\Delta^2/(2\mu + 2\Delta/3)).$$

(See for example [JLR00, Theorems 2.1 and 2.8].) To establish (10) we take $(1 + \varepsilon)n$ i.i.d. Bernoullis with $p_i = (1 - \delta)^2$. For any fixed δ in (10) this immediately gives probability $\exp(-\Theta(n^{6/5}/n))$, and the sum over the $\Theta(n)$ possible values of δ can be subsumed into the exponential.

In the main we will therefore assume that

$$(12) \quad m_2(\delta n) \leq n(1 + \varepsilon)(1 - \delta)^2 + n^{3/5},$$

and deal with the failure case only at the end.

We will also need two simple distributional inequalities. First, a Bernoulli random variable is stochastically dominated by a similar Poisson random variable,

$$\text{Be}(p) \preceq \text{Po}(-\ln(1 - p)),$$

as they give equal probability to 0, and the Bernoulli's remaining probability is entirely on 1 whereas the Poisson's is on 1 and larger values. (Here we have written $\text{Be}(p)$ and $\text{Po}(-\ln(1 - p))$ where we really mean random variables with those distributions; we shall continue this practice where convenient.) Summing n independent copies of such random variables shows that a binomial is dominated by a similar Poisson,

$$B(n, p) \preceq \text{Po}(-n \ln(1 - p)).$$

In particular, for any $a, b = \Theta(1)$,

$$(13) \quad B(an, b/n) \preceq \text{Po}(-an \ln(1 - b/n)) = \text{Po}(ab + O(1/n))).$$

We also recall that the exponential moments of a Poisson random variable are

$$(14) \quad \mathbb{E} z^{\text{Po}(d)} = \exp((z - 1)d).$$

We now analyze the unit-clause algorithm in Phases I and II.

5.2.2 Phase I

During Phase I, assuming (12), at times $t = \delta n$,

$$m_2(t) = n(1 + \varepsilon)(1 - \delta)^2 + O(n^{3/5}) \leq n(1 + 1.01\varepsilon - 2\delta),$$

using $\varepsilon \geq n^{-1/3}$. Meanwhile the number of unset variables is $m(t) = n(1 - \delta)$, so in particular,

$$(15) \quad m_2(t)/m(t) \leq 1 + 1.05\varepsilon.$$

With the random variables below all independent, the unit-clause algorithm gives

$$\begin{aligned} & m_1(t) - m_1(t-1) \\ &= -1 + \mathbf{1}(m_1(t-1) = 0) - B(m_1(t-1), 1/(2m(t-1))) \\ &\quad + B(m_2(t-1), 1/(m(t-1))) \\ &\leq -1 + \mathbf{1}(m_1(t-1) = 0) + B(m_2(t-1), 1/(m(t-1))) \\ &\leq -1 + \mathbf{1}(m_1(t-1) = 0) + \text{Po}(1 + 1.1\varepsilon), \end{aligned}$$

where the last inequality uses (13), (15), and $0.1\varepsilon \gg 1/n$.

It is easy to see that, starting from $m_1(0) = m'_1(0)$, if $X(t) \leq Y(t)$ for all t , if

$$\begin{aligned} m_1(t) - m_1(t-1) &= \mathbf{1}(m_1(t-1) = 0) + X(t-1) \text{ and} \\ m'_1(t) - m'_1(t-1) &= \mathbf{1}(m'_1(t-1) = 0) + Y(t-1), \end{aligned}$$

then for all t ,

$$m_1(t) \leq m'_1(t).$$

(An easy proof is inductive. The $\mathbf{1}(\cdot)$ term may contribute to m_1 and not to m'_1 if $m_1 < m'_1$, but in that case, the inequality still holds.) In a similar setup but with $X(t) \preceq Y(t)$, coupling shows that $m_1(t) \preceq m'_1(t)$.

Thus $m_1(t) \preceq m'_1(t)$ where $m'_1(0) = 0$ and

$$m'_1(t) - m'_1(t-1) = -1 + \mathbf{1}(m'_1(t-1) = 0) + \text{Po}(1 + 1.1\varepsilon).$$

Now, let $U(t)$ be a random walk with $U(0) = 0$ and independent increments

$$(16) \quad U(t) - U(t-1) = -1 + \text{Po}(1 + 1.1\varepsilon),$$

and let $V(t)$ count the “record minima” of U , so $V(0) = 0$ and $V(t) = V(t-1)$ except that if $U(t) < \min_{\tau < t} U(\tau)$, then $V(t) = V(t-1) + 1$. Observe that

$$(17) \quad m_1(t) \preceq m'_1(t) = U(t) + V(t-1).$$

($V(t)$ precisely takes care of the $\mathbf{1}(\cdot)$ terms.)

At this point, we have reduced the behavior of the number of unit clauses $m_1(t)$ to properties of a simple Poisson-incremented random walk.

Renewal process V

We first dispense with V , by showing that

$$(18) \quad V(\infty) \doteq \sup_{t \geq 0} V(t) \preceq G(2\varepsilon),$$

where $G(p)$ indicates a geometric random variable with parameter p . Starting from any time t_0 at which $U(t_0)$ is a record minimum (at which $V(t_0) = V(t_0-1) + 1$), define $U'(\tau) = U(t_0 + \tau) - U(t_0) + 1$. Observe that $U'(0) = 1$, and the first time τ for which $U(\tau) = 0$ gives the next time $t_0 + \tau$ for which $V(t_0 + \tau) = 1$. Thus the number of “restarts” of the process U' is $V(\infty)$.

U' may be viewed as a Galton-Watson branching process observed each time an individual gives birth (adding $\text{Po}(\cdot)$ offspring to the population) and itself dies (adding -1). As a super-critical Galton-Watson branching process, U' has a positive probability of non-extinction, and thus the number of restarts (following extinctions) is geometrically distributed.

Quantitatively, the extinction probability of a Galton-Watson process with X offspring (the probability the process never hits 0) is well known to be the unique root $p \in [0, 1)$ of

$$(19) \quad p = \mathbb{E}(p^X).$$

(See for example [Dur96, pp. 247–248].) Also, for any p such that $p > \mathbb{E}(p^X)$, the probability of non-extinction exceeds $1 - p$. In this case, recalling (16) and (14), we seek p such that

$$p > \mathbb{E}(p^X) = \exp((p-1)(1+1.1\varepsilon))$$

or equivalently, with $q = 1 - p$,

$$\ln(1-q) > -q(1+1.1\varepsilon).$$

Taking a Taylor expansion around $q = 0$ and cancelling like terms, it suffices to ensure that $\frac{1}{2}q + \frac{1}{3}q^2 + \dots < 1.1\varepsilon$, and $q = 2\varepsilon$ suffices (for all $\varepsilon < 0.37$, let alone the $\varepsilon = \Theta(n^{-1/3})$ of interest).

Thus U' has non-extinction probability at least 2ε , verifying (18).

Random walk U

We now analyze the random walk U to show that for any $0 < \varepsilon \leq 0.02$ and $0 < \alpha \leq 0.06$ (our principal realm of interest will be $\varepsilon, \alpha = \Theta(n^{-1/3})$), for any time t ,

$$(20) \quad \Pr \left(\max_{0 \leq \tau \leq t} U(\tau) \geq \mathbb{E}U(t) + \alpha t \right) \leq \exp(-t\alpha^2/2.1).$$

Observe that $U(t)$ is a submartingale, and for any $\beta > 0$ (by convexity of $\exp(\beta u)$), $\exp(\beta U(t))$ is a non-negative submartingale. It follows from Doob's submartingale inequality (see [Dur96]) that

$$\begin{aligned} (21) \quad & \Pr \left(\max_{0 \leq \tau \leq t} U(\tau) \geq \mathbb{E}U(t) + \alpha t \right) \\ &= \Pr \left(\max_{0 \leq \tau \leq t} \exp(\beta U(\tau)) \geq \exp(\beta(\mathbb{E}U(t) + \alpha t)) \right) \\ &\leq \frac{\mathbb{E}(\exp(\beta U(t)))}{\exp(\beta(\mathbb{E}U(t) + \alpha t))}. \end{aligned}$$

Trivially,

$$(22) \quad \mathbb{E}U(t) = -t + (1 + 1.1\varepsilon)t = 1.1\varepsilon t,$$

and, by (14),

$$\begin{aligned} \mathbb{E}(\exp(\beta U(t))) &= \exp(-\beta t + \beta \text{Po}((1 + 1.1\varepsilon)t)) \\ &= \exp(-\beta t) \exp((e^\beta - 1)(1 + 1.1\varepsilon)t), \end{aligned}$$

so (21) is

$$(23) \quad \exp(-t[\beta - (1 + 1.1\varepsilon)(e^\beta - 1) + \beta(1.1\varepsilon + \alpha)]).$$

We are free to choose $\beta > 0$ as we like, so to minimize (23) we maximize the innermost quantity. Setting its derivative equal to 0 yields $1 - (1 + 1.1\varepsilon)e^\beta + 1.1\varepsilon + \alpha = 0$ or $\beta = \ln(1 + \alpha/(1 + 1.1\varepsilon))$, but we will simply take $\beta = \alpha$. Then (eschewing asymptotes in favor of absolute bounds), for $\varepsilon < 0.02$ and $\alpha < 0.06$ (let alone the regime $\varepsilon, \alpha = \Theta(n^{-1/3})$ of interest), (23) is

$$\leq \exp(-t\alpha^2/2.1),$$

proving (20).

Parameter substitution

In the case of interest, $1 \leq \lambda \ll n^{1/3}$, $\varepsilon = \lambda n^{-1/3} \ll 1$, and $t = 2\varepsilon n = 2\lambda n^{2/3}$. Here,

$$\mathbb{E}m_1(t) \leq \mathbb{E}U(t) + \mathbb{E}V(\infty) \leq 1.1\varepsilon t + \frac{1}{2\varepsilon} \leq 2.2\lambda^2 n^{1/3} + n^{1/3} \leq 3.2\lambda^2 n^{1/3}.$$

Substituting $\alpha = \alpha'/\sqrt{t}$ into (20) (which is then valid up to $\alpha' = n^{1/4} = o(n^{1/3})$)

$$(24) \quad \Pr\left(\max_{\tau \leq t} U(\tau) \geq \mathbb{E}U(t) + \alpha'\sqrt{2\lambda}n^{1/3}\right) \leq \exp(-\alpha'^2/2.1),$$

so the tails of $U(t)$ fall off exponentially with a “half-life” smaller than the bound on the mean (as $\lambda \geq 1$ implies $\sqrt{2\lambda} < 2.2\lambda^2$). $V(\infty)$ has an expectation which is at most comparable, and (as a geometric random variable) again falls off exponentially with half-life comparable to its mean.

It follows that

$$\mathbb{E} \sum_{\tau=1}^{2\varepsilon n} m_1(\tau) \leq (2\varepsilon n)\mathbb{E}(U(2\varepsilon n) + V(\infty)) \leq 6.4\lambda^3 n,$$

and, for $\alpha' \leq n^{1/4}$,

$$(25) \quad \Pr\left(\max_{\tau \leq 2\varepsilon n} m_1(\tau) \geq \alpha' 3.2\lambda^2 n^{1/3}\right) = \exp(-\Omega(\alpha')) \text{ and}$$

$$(26) \quad \Pr\left(\sum_{\tau=1}^{2\varepsilon n} m_1(\tau) \geq \alpha' 6.4\lambda^3 n\right) = \exp(-\Omega(\alpha')).$$

The probability of a deviation with $\alpha' > n^{1/4}$ is $\exp(-\Omega(n^{1/2}))$, and will be dealt with as a “failure probability” at the end.

5.2.3 Phase II

The analysis of this phase largely parallels the previous one.

Assuming (12), at times $t = \delta n$, $m_2(t)/m(t)$ is roughly $(1 + \varepsilon)(1 - \delta)$, and in particular, since in Phase II by definition $\delta \geq 2\varepsilon$,

$$(27) \quad m_2(t)/m(t) \leq 1 - 0.95\varepsilon.$$

Since Phase II ends as soon as $m_1(t) = 0$, there is no $+1(\cdot)$ term to worry about, so assuming (27),

$$\begin{aligned} m_1(t) - m_1(t-1) &= -1 - B(m_1(t-1), 1/(2m(t-1))) + B(m_2(t-1), 1/(m(t-1))) \\ &\leq -1 + \text{Po}(1 - 0.9\varepsilon). \end{aligned}$$

By the same argument as for Phase I, then,

$$m_1(2\varepsilon n + t) \preceq m_1(2\varepsilon n) + W(t)$$

where $W(t)$ is a random walk with $W(0) = 0$ and independent increments $-1 + \text{Po}(1 - 0.9\varepsilon)$.

Note that $W(\varepsilon n) = W(\lambda n^{2/3})$ has mean and standard deviation both $\Theta(n^{1/3})$, so for multiples of this time, W is exponentially sure to achieve at least half its (negative) expectation; we now quantify this. At time $\alpha\varepsilon n$,

$$\begin{aligned} \Pr(W(\alpha\varepsilon n) > -\frac{1}{2}\alpha\varepsilon^2 n) &= \Pr(\text{Po}(\alpha\varepsilon n(1 - .9\varepsilon)) > \mathbb{E}(\text{Po}(\cdot)) + 0.4\alpha\varepsilon^2 n) \\ &\leq \exp\left(-\frac{(0.4\alpha\varepsilon^2 n)^2}{\alpha\varepsilon n(1 - 0.9\varepsilon) + 0.4\alpha\varepsilon^2 n}\right) \end{aligned}$$

since the Chernoff bound (11) applies as well to the Poisson. Substituting $\varepsilon = \lambda n^{-1/3}$, the denominator's first term, of order $\Theta(\alpha n^{2/3})$, dominates the second, of order $\Theta(\alpha n^{1/3})$, giving

$$(28) \quad \Pr(W(\alpha\varepsilon n) > -\frac{1}{2}\alpha\varepsilon^2 n) \leq \exp(-0.4^2 \alpha \lambda^3).$$

Then, conditionally on Phase I ending at $m_1(2\varepsilon n) = \alpha_1 \lambda^2 n^{1/3}$ (see (25)), for any $\alpha > 2\alpha_1$, (28) implies that Phase II ends by time $2\varepsilon n + \alpha\varepsilon n$, with probability exponential in α .

Furthermore, over Phase II, $m_1(t)$ is unlikely ever to increase much over its initial value. An argument along the lines used in the context of equation (19) could be constructed to show that $\max_{t \geq 0} W(t)$ is exponentially sure to be quite small, but as there are some technical complications, we take a simple, wasteful approach. Observe that

$$W(t) \preceq X(t)$$

where $X(0) = W(0) = 0$ and $X(t)$ has independent increments $-1 + \text{Po}(1 + \varepsilon)$. This wild over-estimation is useful because X (unlike W) is a sub-martingale, to which we apply Doob's inequality. Just as in sections 5.2.2

and 5.2.2, over a time interval $\alpha\varepsilon n$, X is exponentially unlikely ever to exceed a multiple of its final expectation, $\mathbb{E}X(\alpha\varepsilon n) = \alpha\varepsilon^2 n = \alpha\lambda^2 n^{1/3}$, and so W and in turn m_1 are at least as unlikely to rise more than this amount above their initial values.

So, conditionally on Phase II starting at $m_1 = \alpha_1\lambda^2 n^{1/3}$, Phase II finishes within additional time $\alpha\lambda^2 n^{1/3}$ with exponentially high probability for all $\alpha \geq 2\alpha_1$, and within that additional time, m_1 is exponentially unlikely (in β) to exceed $(\alpha_1 + \beta\alpha)\lambda^2 n^{1/3}$. It follows that if Phase II ends at time $2\varepsilon n + t$,

$$(29) \quad \Pr\left(\sum_{\tau=1}^t m_1(2\varepsilon n + \tau) \geq \text{const } \beta' a_1^2 \lambda^3 n\right) \leq \exp(-\beta').$$

5.2.4 Phases I, II and III

We have argued that over Phases I and II the number of unit clauses $m_1(t)$ is exponentially unlikely ever to exceed a multiple of $\varepsilon^2 n = \lambda^2 n^{1/3}$, and that Phase II is exponentially unlikely to end after a multiple of time $\varepsilon n = \lambda n^{2/3}$, to prove, in (26) and (29), that the summed number of unit clauses $M_1 = \sum_{\tau} m_1(\tau)$ (summed over times τ from 0 to the end of phase II), is exponentially unlikely to exceed a multiple of $\lambda^3 n$:

$$\Pr(M_1 \geq \text{const } \beta \lambda^3 n) \leq \exp(-\beta).$$

By definition of the unit-clause algorithm, at each stage the literals forming the unit clauses are drawn independently at random with replacement from among the literals not yet set, and so the number of unit clauses dissatisfied at each step t is

$$(30) \quad B(m_1(t), 1/(2(n-t)))$$

(where $m_1(t)$ is itself a random variable). With probability $1 - \exp(-\Theta(n^{1/4}))$ these phases end long before time $t = n/3$, so (30) is $\leq \text{Po}(0.8m_1(t)/n)$, and by independence of the random variables in (30) (each conditioned on $m_1(t)$) for different times t , the total number of unit clauses dissatisfied in phases I and II is dominated by $\text{Po}(0.8M_1/n)$.

Since $\mathbb{E}M_1 = O(\lambda^3 n)$, the Poisson's expectation is $O(\lambda^3)$, and the number of X of unit clauses unsatisfied over these phases also has $\mathbb{E}X = O(\lambda^3)$; this confirms (for Phases I and II) *one* assertion of Theorem 6). Fixing $\beta = 1$, there is at least constant probability that $M_1 \leq \text{const } \lambda^3 n$ and so the probability that *no* unit clause is dissatisfied is $\Pr(X = 0) \geq \exp(-O(\lambda^3))$, a *second* assertion of the theorem. Since both M_1 and $\text{Po}(M_1/n)$ have exponential tails, so does X — $\Pr(X \geq \beta \text{ const } \lambda^3) \leq \exp(-\beta)$ — a *third*

assertion of Theorem 6. We now argue that Phase III leaves all these properties intact.

By construction, at the conclusion of Phases I and II the remaining formula is uniformly random, still on $n(1 - o(1))$ variables, but now with density $\leq 1 - \varepsilon \leq 1 - n^{-1/3}$. For Phase III we simply argue that, by the previously proved case $\lambda \leq -1$ of this Theorem, such a formula can be satisfied but for $\leq \text{const } \beta$ clauses, with probability $\geq 1 - \exp(-\beta)$. This concludes the proof of the case $\lambda > 1$ of Theorem 6. \square

5.2.5 Remarks

A corresponding *lower bound*, on the number of clauses that must be violated, cannot be found by the same techniques, since there is no guarantee that the unit-clause algorithm is doing the best possible. One alternative is to analyze the pure-literal rule, which is guaranteed to make no “mistakes” as long as it runs, then use other methods to analyze the remaining “core” formula; we understand that this analysis has been done successfully (and independently) by Kim [Kim]. Another approach might be to extend the “bicycles” analysis of Theorem 3 (or the $\lambda \leq -1$ case of Theorem 6) to the case $\varepsilon > 0$ (particularly, $\varepsilon = \lambda n^{-1/3}$ and $\lambda > 1$), but this seems not to be easy.

We remark that, no matter the particular approach pursued, verifying that the number of clauses that must be dissatisfied is $\Omega(\lambda^3)$ seems to lead back, in intuition and in proof techniques, to the fact that in a $G_{n,p}$ random graph with average degree $np = 1 + \lambda n^{-1/3}$, there is likely to be a giant component whose “kernel” is a random cubic graph on $\Theta(\lambda^3)$ vertices [JLR00, p. 123].

6 Random MAX k-SAT and MAX CSP

In this section we present some general facts and conjectures about MAX k -SAT and MAX CSP, and generalize the 2-SAT high-density results.

6.1 Concentration and limits

It is known that random k -SAT has a sharp threshold: that is, there exists a threshold function $c(n)$ such that for any $\varepsilon > 0$, as $n \rightarrow \infty$, a random formula on n variables with $(c(n) - \varepsilon)n$ clauses is a.s. satisfiable, while one with $(c(n) + \varepsilon)n$ clauses is a.s. unsatisfiable [Fri99]. To prove an analogous result for random MAX k -SAT is much easier; this was first done by [BFU93].

We will employ a “bounded difference” inequality; specifically, a generalization of Azuma’s inequality in a form due to McDiarmid [McD89] (see also Bollobás [Bol88]).

Theorem 10 (Azuma) *Let X_1, \dots, X_n be independent random variables, with X_k taking values in a set A_k for each k . Suppose that the (measurable) function $f : \prod A_k \rightarrow \mathbb{R}$ satisfies $|f(x) - f(x')| \leq c_k$ whenever the vectors x and x' differ only in the k ’th coordinate. Let Y be the random variable $f(X_1, \dots, X_n)$. Then for any $\lambda > 0$, $\mathbb{P}[|Y - \mathbb{E}Y| \geq \lambda] \leq 2 \exp(-2\lambda^2 / \sum c_k^2)$.*

Let $F_k(n, m)$ be a random k -SAT formula on n variables with m clauses, and let $f_k(n, m) = \mathbb{E}(\max F_k)$; we may omit the subscripts k .

Theorem 11 ([BFU93]) *For all k , n , c , and λ , $\mathbb{P}(|\max F_k(n, cn) - f_k(n, cn)| > \lambda) < 2 \exp(-2\lambda^2/(cn))$.*

Proof. Let X_i represent the i th clause in F . Replacing X_i with an arbitrary clause cannot change $\max F$ by more than 1. The result follows from Azuma’s inequality. \square

The theorem’s statement that for any c and large n , $F(n, cn)/(cn)$ has some almost-sure almost-exact value, is reminiscent of Friedgut’s theorem (Theorem 2) that (loosely interpreted) says that for large n and any c away from the threshold, $\Pr(F(n, cn) \text{ is satisfiable})$ is almost exactly either 0 or 1. In our case, the target value $f(n, cn)/(cn)$ is unknown and it is unknown whether it has a limit in n , and in Friedgut’s case, again, it is unknown for which c the probability is near 0 and for which it is near 1, and whether the threshold value of c (and the distribution function) has a limit in n . To conjecture that $f(n, cn)/(cn)$ tends to a limit in n is in this sense analogous to the “satisfiability threshold conjecture”.

Conjecture 12 (MAX SAT limiting function conjecture) *For every k , for every constant $c > 0$, as $n \rightarrow \infty$, $f_k(n, cn)/n$ converges to a limit.*

The conjecture may equally well be extended to arbitrary CSPs, yet is open even for MAX 2-SAT.

If $f_k(n, cn)/(cn)$ were monotone in n , the conjecture’s truth would follow. Of course we do not know this, but can prove monotonicity in c : that as the number of clauses increases, the expected fraction of clauses that can be satisfied can only decrease.

Remark 13 *For any k and n , $f_k(n, m)/m$ is a non-increasing function of m .*

Proof. In a uniform random instance of $F_k(n, m)$, let the maximum number of satisfiable clauses be J , so that $\mathbb{E}(J) = f(n, m)$. By deleting single clauses, we obtain m uniform random instances F of $F(n, m - 1)$. Of these, $m - J$ each have $\max F = J$, while the remaining J each have $\max F \in \{J - 1, J\}$. The average of these m values is at least $\frac{(m-J)(J)+(J)(J-1)}{m} = \frac{J(m-1)}{m}$. Taking expectations, we find $\frac{f(n, m-1)}{m-1} \geq \frac{1}{m-1} \times \mathbb{E}\left(\frac{J(m-1)}{m}\right) = \mathbb{E}\left(\frac{J}{m}\right) = \frac{f(n, m)}{m}$, as desired. \square

Finally, we expect a connection between the MAX SAT limiting function conjecture (Conjecture 12) above and the usual satisfiability threshold conjecture (Conjecture 1). We formalize this in the following conjecture.

Conjecture 14 *For any $c < 0$, $\lim_{n \rightarrow \infty} f(n, cn)/(cn) = 1$ if and only if $\lim_{n \rightarrow \infty} \Pr(F(n, cn) \text{ is satisfiable}) = 1$.*

One aspect of this is easily resolved. If $\limsup f(n, cn)/(cn) < 1$, say $1 - \delta$, then on average $c\delta n$ clauses per formula go unsatisfied, at least a δ fraction of all formulas must be unsatisfiable, and so $\limsup \Pr(F(n, cn) \text{ is satisfiable}) < 1$. But nothing more seems obvious.

6.2 High-density MAX k-SAT and MAX CSP

In this section we extend Theorem 4.

Theorem 15 *For all k , for all c sufficiently large, $(\frac{2^k-1}{2^k}c + \frac{2}{k+1}\sqrt{\frac{ck}{\pi 2^k}} - O(1))n \lesssim f_k(n, cn) \lesssim (\frac{2^k-1}{2^k}c + \sqrt{c}\sqrt{\frac{(2^k-1)\ln 2}{2^{2k-1}}})n$.*

Note that the leading terms are equal, and the second-order terms equal to within $\text{const} \cdot \sqrt{k}$.

Proof. Upper bound. The proof is very similar to that of Theorem 4. Using the first-moment method, we have:

$$\begin{aligned} P &= \mathbb{P}(\exists \text{ satisfiable } F') \\ &\leq 2^n \sum_{l=0}^{rcn} \binom{cn}{l} \left(\frac{2^k-1}{2^k}\right)^{cn-l} \left(\frac{1}{2^k}\right)^l. \end{aligned}$$

For $r < \frac{1}{2^k}$ the sum is dominated by the last term, and so we fix $l = rcn$. Using (1), taking logarithms, and finally substituting $r = \frac{1}{2^k} - \varepsilon$, we have

$$\frac{1}{cn} \ln P \simeq \frac{\ln(2)}{c} - \left(\frac{2^{2k-1}}{2^k-1}\right) \varepsilon^2 + O(\varepsilon^3).$$

Thus for $r < 1/(2^k) - \sqrt{\frac{(2^k-1)\ln 2}{c2^{2k-1}}}$, $P \rightarrow 0$ as $n \rightarrow \infty$.

Lower bound. Set the variables sequentially. Set variables $X_1, X_2, \dots, X_{\ell-1}$ randomly, and then for each $\ell \leq i \leq n$, enumerate those clauses involving only X_i and some subset of $\{X_1, X_2, \dots, X_{\ell-1}\}$ (that is, unit clauses). The expected number of such clauses is about

$$cnk\left(\frac{1}{n}\right)\left(\frac{\ell-1}{n}\right)^{k-1} = ck\left(\frac{\ell-1}{n}\right)^{k-1},$$

and if we count only those left unsatisfied by their previous $k-1$ variables, the expected number becomes

$$h_\ell = \frac{ck}{2^{k-1}}\left(\frac{\ell-1}{n}\right)^{k-1}.$$

(Here we incur a minor error by sampling with replacement instead of without; $(\frac{\ell-1}{n})^{k-1}$ should really be $\prod_{1 \leq h \leq k-1} \left(\frac{\ell-h}{n-h}\right)$.) More precisely, the number of such clauses enjoys a Poisson distribution with mean $h\ell$. Set the value of X_i to maximize the number of such clauses satisfied; as before, this number is about $\frac{1}{2}h_\ell + \frac{1}{2}\sqrt{\frac{2}{\pi}h_\ell} + O(1)$. The advantage over purely random guessing is

$$\sqrt{\frac{1}{2\pi}h_\ell} + O(1) = \sqrt{\frac{ck}{2\pi 2^{k-1}}\left(1 - \frac{\ell-1}{n}\right)^{k-1}} + O(1).$$

Sum over $i = \ell, \dots, n$ to obtain an advantage of

$$\sqrt{\frac{ck}{\pi 2^k} \frac{2n}{k+1}} + O(n).$$

□

Still more generally, we may consider a CSP (constraint satisfaction problem). Let g be a k -ary “constraint” function, $g : \{0, 1\}^k \rightarrow \{0, 1\}$. A random formula $F_g(n, m)$ over g is defined by m clauses, each chosen uniformly at random (with replacement) from the $2^k n(n-1) \cdots (n-k+1)$ possible clauses defined by an ordered k -tuple of distinct variables each appearing positively or negated. (Formally, a clause consists of a k -tuple (i_1, \dots, i_k) of distinct values in $[n]$, specifying the variables, and a binary k -vector $(\sigma_1, \dots, \sigma_k)$, specifying their signs.) A clause with variables (signed variables) X_1, \dots, X_k is satisfied if $g(X_1, \dots, X_k) = 1$. (Formally, an assignment x_1, \dots, x_n of the full set of variables X_1, \dots, X_n satisfies a clause as above if $g(x_{i_1} \oplus \sigma_1, \dots, x_{i_n} \oplus \sigma_n) = 1$, where “ \oplus ” denotes XOR, or addition

modulo 2.) As ever, such a formula F is satisfiable if there exists an assignment of the variables satisfying all the clauses; and $\max F$ is the maximum, over all assignments, of the number of clauses satisfied.

Generally a CSP may be based on a finite family of constraint functions, of “arities” bounded by k , but for notational convenience we limit ourselves to a single function.

Let a k -ary clause function g be given, with $\mathbb{E}(g(X)) = p$ over random inputs. Define $P = \min\{p, 1-p\}$ and $Q = 1 - P$. Let $F_g(n, m)$ be a random formula over g on n variables, with m clauses, and let $f_g(n, m) = \mathbb{E}(\max F)$.

Theorem 16 *Given an arity k and a constraint function g , for all c sufficiently large, $(pc + \sqrt{PQ^2c/k})n \lesssim f_g(n, m) \lesssim (pc + \sqrt{2PQ \ln(2)c})n$.*

The proof follows that of Theorem 15, and is omitted.

7 Online random MAX 2-SAT

In this section, we discuss online versions of the MAX 2-SAT problem. [BF01, BFW02] consider an online version of MAX giant-free spanning subgraph, in which random edges e_i are given one by one, and we must accept or reject e_i based on the previous edges e_1, \dots, e_{i-1} , with the goal of accepting as many edges as possible without creating a giant component.

There are two natural online interpretations of random MAX 2-SAT. In both, we are told in advance the total number of variables n and clauses m ; also, in both, clauses c_i are presented one by one, and we must choose “on line” whether to accept or reject c_i based on the previously seen clauses c_1, \dots, c_{i-1} . When we accept a clause we are guaranteeing to satisfy it; when we reject a clause we are free to satisfy or dissatisfy it. Our goal is to maximize the number of clauses accepted.

In our first interpretation of online MAX 2-SAT, ONLINE I, when we accept a clause, we are also required to satisfy it immediately, by setting at least one of its literals True; once a variable is set, it may never be changed. The second interpretation, ONLINE II, is more generous: the variables’ assignments may be decided after the last clause is presented. Let $f_{\text{O-I}}(n, m)$ be the expected number of clauses accepted by an optimal algorithm for ONLINE I, and $f_{\text{O-II}}(n, m)$ that for ONLINE II. Clearly, $\frac{3}{4}m \leq f_{\text{O-I}}(n, m) \leq f_{\text{O-II}}(n, m) \leq f(n, m)$. Here we present a “lazy” algorithm applicable to both $f_{\text{O-I}}(n, cn)$ and $f_{\text{O-II}}(n, cn)$. ONLINE-LAZY begins with no variables “set”. On presentation of a clause, ONLINE-LAZY rejects it only if it must, and otherwise does the least it can to accept it. Specifically, on

presentation of clause c_i , which without loss of generality we may consider to be $(X \vee Y)$, it takes the following action. If $X = \text{True}$ or $Y = \text{True}$, accept c_i . If $X = \text{False}$ and $Y = \text{False}$, reject c_i . If $X = \text{False}$ and Y is unset (or vice-versa), set $Y = \text{True}$ (resp. $X = \text{True}$) and accept c_i . If X and Y are both unset, arbitrarily choose one, set it True, and accept c_i .

Theorem 17 *For any fixed c , ONLINE-LAZY is the unique (up to its arbitrary choice) optimal algorithm for ONLINE I, and $f_{\text{O-I}}(n, cn) \simeq (\frac{3}{4}c + (1 - e^{-c})/4 + (1 - e^{-c})^2/8)n \geq (\frac{3}{4}c + \frac{3}{8})n$.*

We note that for $c = 1$, $f_{\text{O-I}}(n, n) \approx 0.957997n$, and for c asymptotically large, $f_{\text{O-I}}(n, cn) \simeq (\frac{3}{4}c + \frac{3}{8})n$.

Proof of optimality. On appearance of a clause c_i , it is clearly best not to set any variable not appearing in c_i , for this merely imposes extra constraints. Similarly, if c_i is already satisfied by one of its literals, then it is best to accept it and to set no additional variables.

The only interesting cases, then, are if c_i is not already satisfied, but one or both of its variables are unset. Again, if both variables are unset, it is best to set at most one of them, and it doesn't matter which one: the "future" performance of an optimal algorithm is solely a (random) function of the number of unset variables and the number of clauses remaining, and these parameters of the future, as well as the number of clauses accepted in the past, are the same whether c_i 's first or second literal is set.

It only remains to show that if c_i is not satisfied by a variable already set, and at least one of its variables is not yet set, then an optimal algorithm must set one of its literals to True. Consider a putatively optimal algorithm Opt which does not do this, so for a literal X in c_i , either Opt sets X to False, or it leaves X unset.

In the case when Opt sets X to False, let a competing algorithm Opt' set X to True, then simulate Opt but reversing the roles of X and \bar{X} in future clauses. "Couple" the distribution of future random clauses seen by Opt and Opt', also by reversing the roles of X and \bar{X} . With this coupling, Opt' accepts exactly the same number of clauses as Opt in the future, but has accepted one additional clause so far (c_i); this contradicts the supposed optimality of Opt.

The slightly less obvious case is when Opt leaves X unset. Again we introduce a competing algorithm Opt', which sets X to True, then simulates Opt until such time as Opt sets X . For inputs where Opt never sets X , Opt' accepts every clause that Opt accepts, as well as the clause c_i , and perhaps additional clauses in which X appears; Opt' is strictly better on

these inputs. For inputs where Opt eventually sets X to True, Opt' goes on simulating Opt, again performing exactly as well on future clauses, and strictly better on past ones. For inputs where at time $j > i$, Opt sets X to False, Opt' may simulate Opt but (as before) with the roles of X and \bar{X} reversed. With the previous coupling, on these inputs, Opt' accepts exactly as many future clauses as Opt, and at least as many in the past (Opt' has accepted c_i and perhaps other clauses rejected by Opt, while Opt has accepted c_j and no other clause rejected by Opt'). So in all three cases, the expected number of clauses accepted by Opt' is at least as many as for Opt, and in the first two cases, which occur with nonzero probability (for example, if no future clause contains X), strictly more; this contradicts the supposed optimality of Opt.

Proof of performance.

Note that clauses causing a variable to be set by ONLINE-LAZY are always satisfied, and those not causing a variable to be set are satisfied with probability 3/4 (if both variables are set) or 1 (if one is set satisfyingly).

If k variables are yet to be set, the probability that a clause has neither variable set is $(k/n)^2$, the probability it has one variable set non-satisfyingly and the other not set is $2 \cdot \frac{1}{2} \cdot ((n-k)/n)(k/n)$, so a random clause falls into one of these cases w.p. k/n . The expected time to set another variable when k are unset is thus n/k . In this period, clauses have (unconditioned) probabilities $(n-k)^2/n^2$ that both variables are set, and $k(n-k)/n^2$ that one is set satisfyingly and the other unset; conditional upon one or other of these being the case (a variable is not set for this clause), the probabilities are $(n-k)/n$ for the first case and k/n for the second, and the clause is satisfied with probabilities 3/4 and 1 in these cases, for average gain $\frac{1}{4}k/n$ over the naive 3/4. The total gain in the number of clauses satisfied in the expected $n/k - 1$ steps before the setting, and the n/k 'th step with the setting, is $(\frac{n}{k} - 1)(\frac{1}{4}k/n) + 1/4 = 1/2 - \frac{1}{4}k/n$. The process goes through $k = n, n-1, \dots, n-I^*$, until the sum of the waiting times exceeds the number of clauses cn . Where $H(i)$ denotes the i 'th harmonic number, for a given I , the *expected* sum of the waiting times is $\sum_{i=0}^I n/(n-i) = n(H(n) - H(n-I-1)) \approx n(\ln(n/(n-I)))$. Solving for this equal to cn gives $n/(n-\hat{I}) = \exp(c)$, or $\hat{I} = n(1 - \exp(-c))$.

What is the variance in the total waiting time W , for $I = \hat{I}$, and where we will allow the total to exceed cn ? Each individual waiting time is geometrically distributed with a mean in the range $\frac{n}{n} = 1$ to $\frac{n}{n(\exp(-c))} = \exp(c)$, all of which are $O(1)$, so W has standard deviation $O(\sqrt{n})$. The amount by which we may have overshot (or fallen short of) the target value cn

is $W - cn$; since each round takes time at least 1, to reach precisely cn it suffices to back off (or add) at most $W - cn$ rounds. That is, $|I^* - \hat{I}| \leq |W - cn|$, which with probability exponentially close to 1 is $O(n^{2/3})$. The expected total number of clauses satisfied over the naive $3/4$ fraction is then $\mathbb{E} \left(\sum_{i=0}^{I^*} (1/2 - \frac{1}{4}(n-i)/n) \right) \simeq \hat{I}/4 + \hat{I}^2/(8n)$. That is, the expected number of clauses satisfied is $\simeq (\frac{3}{4}c + (1 - e^{-c})/4 + (1 - e^{-c})^2/8)n$. \square

Note that ONLINE-LAZY does not, in fact, need to know the number of clauses in advance!

A variant of ONLINE I is that if we accept a clause we must set *both* its variables. In this case, similar arguments show that an optimal algorithm simply sets each new literal True.

We know essentially nothing about ONLINE II. To obtain improved bounds, or, ideally, to identify a provably optimal algorithm, are interesting open problems.

8 Random MAX CUT

8.1 Motivation

One source of motivation for our work was, as mentioned in the introduction, that although *random* constraint satisfaction problems (CSPs) and *max* CSPs are well studied, random MAX CSPs seem not to have been. However, we had a second, particular source of motivation, in recent work on “avoiding a giant component” in a random graph.

Think of MAX SAT as the problem of, given a formula, to select as many clauses as possible so that the subformula of selected edges is satisfiable. An analogous problem is, given a graph, to select as many edges as possible so that the subgraph of selected edges has no giant component (suitably defined).

The latter problem was posed in a slightly different form by Achlioptas, who asked how many random edge *pairs* could be given, such that by selecting one edge from each pair, a giant component could be avoided. Bohman and Frieze showed in [BF01] that a giant component can be avoided with $0.55n$ edge pairs (where a random selection of one edge from each pair would almost surely generate a giant component). Bohman, Frieze, and Wormald [BFW02] considered the problem without Achlioptas’s original “pairing” aspect: how many edges may a random graph have, so that some subgraph with $1/2$ the edges has no giant component. They show that this is true up

to about $1.958n$ edges but not beyond (where the precise threshold satisfies a transcendental equation). Without the pairing aspect, there is no longer anything special about $1/2$, though, and [BFW02] is easily extended to answer the question: for a random graph $G(n, cn)$, how many edges $f(n, cn)$ may be retained while avoiding a giant component. This is precisely the same sort of question we considered for SAT, and was in our minds when we began this work.

It is tempting to imagine a particular connection between the two questions, because of a well known connection between the unsatisfiability of a random 2-SAT formula and the existence of a giant component in a random graph, most easily explained in terms of branching processes. For a 2-SAT formula F , consider a branching process on literals, where a literal X has offspring including Y if F includes a clause $\{\bar{X}, Y\}$ (and if Y was not the parent of X). (The process models the fact that if X is set true, Y must also be set true to satisfy F). Although additional work is needed to prove it, a random 2-SAT formula is satisfiable with high probability if this branching process is subcritical (if each X has an expected number of offspring < 1) and unsatisfiable w.h.p. if it is supercritical. For a random graph G , consider a branching process on vertices, where a vertex v has offspring including w if G has an edge $\{v, w\}$ (and if w was not the parent of v). Here, w.h.p. G has no giant component if the process is subcritical, and w.h.p. has one if it is supercritical. These intuitively explain the phase-transition thresholds of cn clauses, $c = 1$, for a random 2-SAT formula, and edge density c/n , $c = 1$, for a random graph.

Despite this connection between unsatisfiability of a random formula, and a giant component in a random graph, the size of a largest giant-free subgraph of a random graph behaves very differently from the size of a largest satisfiable subformula of a random formula. Specifically, for large clause density c , there is a satisfiable subformula preserving an expected constant fraction ($3/4$ ths) of the clauses, while for a random graph with cn edges, the largest giant-free subgraph has only about n edges, a $1/c$ fraction. This can be read off from Theorem 18, or argued more simply: if G had a giant-free subgraph H with linearly more than n edges, H (and thus G) would have to have a linear-size dense component, but a random sparse graph has no linear-size dense component.

Define $f_{\text{nogiant}}(n, m) \doteq \mathbb{E}(\text{MAX giant-free}(G(n, m)))$.

Theorem 18 *With $t = t(c) < 1$ defined by $te^{-t} = 2ce^{-2c}$, $f_{\text{nogiant}}(n, cn) = rcn$ when $\frac{t^2}{4c} + 1 = \frac{t}{2c} + cr$.*

The theorem is proved as in [BFW02] (modifying their Lemma 1 to allow

values $c > 2$ by replacing a $(\log n)/6$ with $(\log n)/(6 \log c)$.

Is there another MAX subgraph problem, then, which does behave like MAX 2-SAT? Going back to the branching process for a random graph — the source of the intuitive connection between the graph and SAT problems — it is also easy to check that w.h.p. a graph has few cycles when the branching process is subcritical, and many cycles when it is supercritical. So perhaps we should consider the size of maximum cycle-free subgraph. But this is by definition a forest, which may have at most $n - 1$ edges, again a $1/c$ fraction, not a fixed constant fraction as for MAX 2-SAT.

In a 2-SAT formula, obstructions to satisfiability come not from cycles of implications $X \implies \dots \implies X$, but only from those with $X \implies \dots \implies \bar{X}$. By a very vague analogy, then, perhaps on the graph side we should seek not a subgraph which is entirely cycle-free, but just one which is free of *odd* cycles: a bipartite subgraph. The size of a largest bipartite subgraph H of G is by definition, and more familiarly, the size of a maximum cut of G . Here, finally, we share with MAX 2-SAT that we may keep a constant fraction of the input structure: for a random graph (indeed any graph) G of size m , $\text{MAX CUT}(G) \geq m/2$, since a random cut achieves this expectation.

8.2 MAX CUT

In addition to the fact that just as a maximum assignment satisfies at least $3/4$ ths the clauses of any formula, a maximum cut cuts at least $1/2$ the edges of a graph, there are other commonalities.

MAX CUT, like MAX 2-SAT, is a constraint satisfaction problem (CSP). With each vertex v we associate a boolean variable representing the partition to which v belongs, and with each edge $\{u, v\}$ we associate a “cut constraint” $(u \oplus v)$, these XOR constraints replacing 2-SAT’s disjunctions. Like decision 2-SAT, the problem of whether a graph is perfectly cuttable (bipartite) is solvable in essentially linear time. In further analogy with MAX 2-SAT, MAX CUT is NP-hard, trivially $\frac{1}{2}$ -approximable, 0.878-approximable [GW95] by semidefinite programming, and not better than 16/17-approximable [TSSW00] in polynomial time, unless P=NP.

The methods we have applied to random MAX 2-SAT are equally applicable to MAX CUT, and yield analogous results. Because it is easier to work with random graphs than random formulas, and more is known about them, our results for MAX CUT are in some respects stronger than those for MAX 2-SAT.

When we work in the $G(n, p)$ model we will take $p = 2c/n$, and in the $G(n, m)$ model, $m = \lfloor cn \rfloor$, so that in both cases the phase transition occurs

at $c = 1/2$. We now state our main results.

8.3 Results

Theorem 19 *For $c = 1/2 - \varepsilon(n)$, with $n^{-1/3} \ll \varepsilon(n) < 1/2$, $f_{\text{cut}}(n, \lfloor cn \rfloor) = \lfloor cn \rfloor - \Theta(\ln(1/\varepsilon)) + \Theta(1)$.*

In particular, for small constants ε this gap of $\Theta(\ln(1/\varepsilon))$ — which for a fixed ε is $\Theta(1)$ — contrasts with the gap of $\Theta(1/n)$ for MAX 2-SAT (Theorem 3). But here too there is a phase transition, in that for $c > 1/2$ the gap jumps to $\Theta(n)$, per Theorem 21.

Theorem 20 *For c large, $\left(\frac{1}{2}c + \sqrt{c} \cdot \sqrt{8/(9\pi)}\right)n \lesssim f_{\text{cut}}(n, cn) \lesssim \left(\frac{1}{2}c + \sqrt{c}\sqrt{\ln(2)/2}\right)n$.*

The values of $\sqrt{8/(9\pi)}$ and $\sqrt{\ln(2)/2}$ are approximately 0.531922 and 0.588704, respectively. The upper bound was previously obtained in [BCP97].

Theorem 21 *For any fixed $\varepsilon > 0$, $(\frac{1}{2} + \varepsilon - \frac{16}{3}(\varepsilon^3))n \lesssim f_{\text{cut}}(n, (1/2 + \varepsilon)n) \lesssim (\frac{1}{2} + \varepsilon - \Omega(\varepsilon^3/\ln(1/\varepsilon)))n$.*

The upper bound's $\varepsilon^3/\ln(1/\varepsilon)$ can probably be replaced by ε^3 , just as we suspect it can be for Theorem 5. This presumption is largely based on the next “scaling window” result.

Theorem 22 *For any function $\varepsilon = \varepsilon(n)$ with $n^{-1/3} \ll \varepsilon(n) \ll 1$, $f_{\text{cut}}(n, (1/2 + \varepsilon)n) = (\frac{1}{2} + \varepsilon - \Theta(\varepsilon^3))n$.*

That the theorem misses out the extremes $\varepsilon = \Theta(n^{-1/3})$ and $\varepsilon = \Theta(1)$ that are perhaps of greater interest than the mid-range is a direct carryover from the standard results on random graphs on which we based our proof is based; it is likely that other established results for random graphs could complete the picture.

Before proceeding, we remark that bipartiteness is of course the same as 2-colorability, and it is sometimes convenient to speak of coloring vertices black or white, rather than placing them in the left or right part of a partition, with properly colored edges (with one black and one white endpoint) corresponding to cut edges; these two ways of speaking are of course mathematically identical.

8.4 Subcritical MAX CUT

THEOREM 19: *Proof.* For notational convenience we work in the $G(n, p)$ model, $G = G(n, (1 - 2\varepsilon)/n)$, but the proof follows identically for the $G(n, m)$ model.

Tree components of G can be cut perfectly; each unicyclic component can be cut for all but 1 edge; and complex components, where more edges must go uncut but which with high probability are absent from G , contribute negligibly. That is, $\mathbb{E}(\#\text{uncut edges}) = (1-o(1))\mathbb{E}(\#\text{cycles in } G)$. Since the number of potential k -cycles is $(n)_k/(2k)$, where $(n)_k = n(n-1)\cdots(n-k+1)$ denotes falling factorial, using $(n)_k = n^k \exp(-k^2/(2n) - O(k/n + k^3/n^2))$ (see [JLR00, eq (5.5)]),

$$\begin{aligned}\mathbb{E}(\#\text{cycles in } G) &= \sum_{k=3}^n \frac{(n)_k}{2k} (c/n)^k \\ &= \sum \frac{1}{2k} c^k \exp(-k^2/(2n) - O(k/n + k^3/n^2)).\end{aligned}$$

Because of the c^k , up to constant factors we need consider the sum only up to $k \leq 1/\varepsilon$ (recalling $c = 1 - 2\varepsilon$), and since $\varepsilon \gg n^{-1/3}$, this makes the entire final exponential term negligibly close to 1. Thus

$$\begin{aligned}\mathbb{E}(\#\text{cycles in } G) &= \Theta(1) \sum_{k=3}^{\infty} c^k / (2k) \\ &= \Theta(1) \left(-\frac{1}{2} \ln(1 - c) \right) - \Theta(1) \\ &= \Theta(1) \ln(1/(2\varepsilon)) - \Theta(1),\end{aligned}$$

where the final $\Theta(1)$ term lies between 0 and $3/2$. \square

8.5 High-density random MAX CUT

THEOREM 20: *Proof.* For the upper bound, we apply a first-moment argument identical to that used in the proof of Theorem 4. The probability that there exists a (maximal) bipartite spanning subgraph of size $\geq (1-r)cn$ is $P \lesssim 2^n \binom{cn}{rcn} (1/2)^{(1-r)cn} (1/2)^{rcn}$, for $\frac{1}{cn} \ln P \lesssim \ln 2/c - r \ln r - (1-r) \ln(1-r) - \ln 2$. Substituting $r = 1/2 - \varepsilon$ gives $\frac{1}{cn} \ln P \lesssim \ln 2/c - 2\varepsilon^2$, so if $\varepsilon > \sqrt{\ln(2)/(2c)}$ then $P \rightarrow 0$.

For the lower bound, color the vertices in random sequence. When xn vertices have been colored, with $x = \Theta(1)$, since c is large, the next vertex is a.s. adjacent to a.e. $2cx$ of the colored ones. In the worst case, the colored vertices are half black and half white; coloring the new vertex oppositely to the majority beats cx (in expectation) by $\mathbb{E}(|B(2cx, 1/2) - cx|) \simeq \mathbb{E}(|N(0, cx)|) = \sqrt{2cx/\pi}$. Integrating over x from 0 to 1 gives $n\sqrt{2c/\pi} \cdot \frac{2}{3}$ more properly colored edges than the naive $\frac{1}{2}cn$. \square

8.6 Low-density random MAX CUT

The following fact follows from small- ε asymptotics of classical random graph results; see, e.g., [Bol98, VII.5, Theorem 17].

Claim 23 *For $\varepsilon > 0$, a random graph $G(n, (1/2 + \varepsilon)n)$ a.s. has a giant component of size $(4\varepsilon + o(\varepsilon))n$.*

Proof. It is well known (see, e.g., [Bol98, VII.5, Theorem 17]) that for an arbitrarily slowly growing function $w(n)$, a.s., the size $L^{(1)}(G)$ of the giant component satisfies $|L^{(1)}(G) - \gamma n| \leq w(n)n^{1/2}$ where $0 < \gamma < 1$ is the unique solution of $e^{-2c\gamma} = 1 - \gamma$. (We have $2c$ where [Bol98] has c because we use cn edges where it uses average degree c .) Take the asymptotic approximation when $c = 1/2 + \varepsilon$. \square

Claim 24 *The probability that a random graph $G(n, (1/2 + \varepsilon)n)$ is bipartite, conditioned on the existence of a component of size $\Theta(\varepsilon n)$ created by the “first” $(1/2 + \varepsilon/2)n$ edges, is $\exp(-\Omega(\varepsilon^3 n))$.*

Proof. If the presumed giant component is not bipartite, we are done. If it is, by connectivity, it has a unique bipartition; let the sizes of the parts be n_1 and n_2 . Each of the remaining $\varepsilon n/2$ edges has both endpoints in the giant component w.p. $\Theta(\varepsilon^2)$, so there are $\Theta(\varepsilon^3 n)$ of these, w.p. $1 - \exp(-\Omega(\varepsilon^3 n))$. The probability that each such edge preserves bipartiteness is $(2n_1 n_2)/(n_1 + n_2)^2 \leq 1/2$; over the $\Theta(\varepsilon^3 n)$ independent edges it is $\exp(-\Omega(\varepsilon^3 n))$. \square

THEOREM 21: Proof. For the upper bound, the first-moment method is applied exactly as in the proof of Theorem 5. We use the preceding Claim, and replace its Ω with an α_0 for definiteness. With $c = (1/2 + \varepsilon)$, then, the probability that deleting any $k \leq rcn$ edges can leave a bipartite subgraph is $P \leq \sum_{k=0}^{rcn} \binom{cn}{k} \exp(-\alpha_0(\varepsilon - k/n)^3)$. This is just as in inequality (2), so here again we conclude that $r \gtrsim \alpha_0 \varepsilon^3 / \ln(1/\varepsilon)$.

The proof of the lower bound is algorithmic, and in direct analogy to that of Theorem 5. Think of a graph edge neither of whose vertices has yet been colored as a “2-clause”, an edge one of whose vertices has been colored as a “unit clause” implying the opposite color for the remaining vertex, an edge whose two vertices have been colored alike as an “unsatisfied clause”, and an edge whose two vertices have been colored oppositely as a “satisfied clause”. Terminate if there are no unit clauses nor 2-clauses. If there are no unit clauses, randomly color a random vertex from a random edge. If there are unit clauses, choose one at random and color its vertex satisfyingly.

Note that when a δ fraction of the vertices have been colored, $(1 - \delta)n$ vertices remain uncolored (unfixed variables), and a.s. a.e. $(1/2 + \varepsilon)n \cdot (1 - \delta)^2$ 2-clauses remain. Each time a unit-clause variable is set, each 2-clause has probability $2/((1 - \delta)n)$ of generating a unit clause

Thus the expected number of 2-clauses becoming unit clauses is $2[(1/2 + \varepsilon)n(1 - \delta)^2]/[(1 - \delta)n] \approx 1 + 2\varepsilon - \delta$, while the number of unit clauses eliminated (satisfied or unsatisfied) is at least 1. Thus the expected increase per step in the number of unit clauses is at most $2\varepsilon - \delta$. As in the proof of Theorem 5, over the first $4\varepsilon n$ steps, the expected number of unit clauses is bounded by an inverted parabola of base $4\varepsilon n$ and height $2\varepsilon^2 n$. Improperly colored edges result only from violated unit clauses, and the expected number of these in the first $4\varepsilon n$ steps is $\leq \frac{2}{3} \cdot 4\varepsilon n \cdot 2\varepsilon^2 n/n \leq \frac{16}{3}\varepsilon^3$. By step $4\varepsilon n$ there are no unit clauses, and the number of 2-clauses divided by the number of unset variables is a.s. a.e. $[(1/2 + \varepsilon)n \cdot (1 - 4\varepsilon)^2]/[n \cdot (1 - 4\varepsilon)] = 1/2 - \varepsilon$. This is a sparse random graph, which by Theorem 19 can be colored to violate just $\Theta(1)$ edges.

In toto, all but $\lesssim (\frac{16}{3}\varepsilon^3 n + \Theta(1))$ edges are properly colored. \square

8.7 Scaling window

The proof of Theorem 22 follows rather easily from standard — but relatively recent, and lovely — facts about the kernel of a random graph. The following summary of the relevant facts, which we present informally, is distilled from [JLR00, Sec. 5.4].

First, if $i \gg n^{2/3}$, then the number of vertices of $G(n, n/2 + i)$ belonging to unicyclic components is asymptotically almost surely $\Theta(n^2/i^2)$. Consider the components of a graph G which are trees, unicyclic, or complex. In the supercritical phase with $n^{-1/3} \ll \varepsilon \ll 1$, a random graph $G(n, (1/2 + \varepsilon)n)$ consists of tree components, unicyclic components, and no complex component other than a single “giant component”. The expected number of vertices in the cycles of the unicyclic components is of order $1/\varepsilon$. The giant

component's 2-core has order $(1 + o(1))8\varepsilon^2n$, and is obtained as a random subdivision of the edges of a "kernel", which is a random cubic graph on $(1 + o(1))\frac{32}{3}\varepsilon^3n$ vertices.

THEOREM 22: *Proof.* We consider which edges of G it may be impossible to cut. Every edge in the tree components of $G = G(n, (1/2 + \varepsilon)n)$ can of course be cut. For each unicyclic component, at most 1 edge must go uncut (if the cycle is odd). By the symmetry rule (see for example [JLR00, Theorem 5.24]), the number of unicyclic components for $G(n, (1/2 + \varepsilon)n)$ is essentially the same as for $G(n, (1/2 - \varepsilon)n)$, which by Theorem 19 is only $O(\ln(1/\varepsilon))$.

The dominant contribution will come from the giant component. Edges which are not in its 2-core can of course all be cut, even after a partition of the 2-core has been decided. Moreover, an optimal partition of the 2-core is essentially decided by a partition of the vertices of the "kernel", which is the 2-core where each path whose internal vertices are all of degree 2 is replaced by a single edge. (See [JLR00, Chap. 5.4] for more on the giant component, its core, and its kernel.) For any cut of the kernel, each 2-core path corresponding to a kernel edge can be partitioned either perfectly or with one edge uncut, depending on the parity of the path's length and whether its endpoints are on the same side or opposite sides of the kernel's cut. Equivalently, a kernel edge whose 2-core path is of odd length imposes a "cut" constraint on its endpoints, while a kernel edge whose 2-core path is of even length imposes an "uncut" constraint on its endpoints; the number of these constraints violated by a cut of the kernel vertices is equal to the number of original cut constraints violated by an optimal extension of the same cut to all the 2-core vertices (and indeed to all the giant-component vertices).

Since each kernel edge is randomly subdivided, on average into $3/(4\varepsilon)$ 2-core edges, the parities of the kernel edges are almost perfectly random (with the probability of either parity approaching $1/2$ as ε approaches 0). For our purposes it suffices that either parity occurs with probability at most some absolute constant $p_0 < 1$, and using this we show that at least some constant fraction β_0 of the approximately $16\varepsilon^3n$ edge constraints must be violated.

Fix a spanning tree T of the kernel K , whose order we will write as N (expecting $N \approx \frac{32}{3}\varepsilon^3n$). Let K subsume not only the graph but also the edge parities, so that it is an instance of the generalized (cut/uncut) MAX CUT problem. If it is possible to violate precisely a fraction $\beta < \beta_0$ of K 's constraints then reversing precisely those constraints gives a perfectly

satisfiable cut/uncut constraint problem instance K' .

Fixing the “side” of any one vertex, the $N - 1$ constraints from the spanning tree T imply the rest of the cut, which must then satisfy the remaining $\frac{1}{2}|N| + 1$ constraints. Viewing the parities of the spanning tree edges as arbitrary, and the remaining edges as independent random variables, the probability that the randomly chosen kernel edges satisfy each of these constraints is at most $p_0^{\frac{1}{2}N+1}$. The number of choices of $t < \beta_0 \frac{3}{2}N$ edges to dissatisfy is $\binom{\frac{3}{2}N}{t}$. We guarantee an exponentially small probability of success by selecting β_0 to satisfy:

$$\begin{aligned} \sum_{t < \beta_0 \frac{3}{2}N} \binom{\frac{3}{2}N}{t} p_0^{\frac{1}{2}N+1} &\ll 1 \\ \frac{3}{2}NH(\beta_0) + \frac{1}{2}N\ln(p_0) &< 0 \\ H(\beta_0) &< \frac{1}{3}\ln(1/p_0), \end{aligned}$$

where H is the entropy function $H(x) = x\ln(x) - (1-x)\ln(1-x)$. In particular, in the case of interest where $\varepsilon \rightarrow 0$, $p_0 \rightarrow 1/2$ and $\beta_0 \rightarrow H^{-1}(1/3) \approx 0.896$. Recapitulating, we must dissatisfy $\beta_0 N$ kernel constraints, $= (32\beta_0/3)\varepsilon^3 n$ constraints of G . The expected $O(\ln(1/\varepsilon))$ uncut edges from unicursal components are negligible by comparison, so in all $\Theta(\varepsilon^3 n)$ edges of G go uncut.

□

9 Conclusions and open problems

We have presented a road map for MAX 2-SAT and MAX CUT in a random setting, establishing that there is a phase transition, and deriving asymptotics below the critical value, for constants slightly above the critical value and in the scaling window around it, and for larger constants.

For constant densities slightly above threshold there is a logarithmic gap between our lower and upper bounds; we need to confirm that the $\ln(1/\varepsilon)$ factors are extraneous. In the other cases, our bounds are only separated by a constant. However, in light of the exact result of [BFW02] for the size of a maximum subgraph which has no giant component, it would be wonderful to get the *exact* asymptotics of $f(n, cn)/(cn)$.

Whether $f(n, cn)/(cn)$ tends to a limit in n (see Conjecture 12) is to our minds a prime open problem in this area, and is not only in some sense

analogous to the satisfiability threshold conjecture, but may also be directly connected with it (see Conjecture 14), another important question.

A question similar in spirit to Conjecture 12 was considered in [Gam02], which defines a certain linear-programming relaxation of MAX 2-SAT. An instance is characterized by its “distance to feasibility” O , with $O(n, cn)$ the corresponding random variable for a random instance. It is shown that for every $c > 0$, $O(n, cn)/(cn)$ almost surely converges to a limit. The result is established using powerful local weak convergence methods [Ald92, Ald01, AS02]. It remains to be seen whether these methods are applicable to random maximum constraint satisfaction problems, including MAX 2-SAT and MAX CUT.

Acknowledgments

Dr Sorkin is very grateful to Svante Janson for pointing out the beautiful properties of the kernel of a random graph and its relevance, and for other helpful technical discussions. Dr Sorkin is also grateful to Mark Jerrum, Martin Dyer, and Peter Winkler for organizing the excellent Newton Institute workshop which fostered these exchanges.

References

- [Ald92] D. Aldous, *Asymptotics in the random assignment problem*, Probab.Th.Rel.Fields (1992), no. 93, 507–534.
- [Ald01] ———, *The $\zeta(2)$ limit in the random assignment problem*, Random Structures and Algorithms (2001), no. 18, 381–418.
- [AM02] Dimitris Achlioptas and Cris Moore, *The asymptotic order of the k -SAT threshold*, Proceedings of the 43rd Annual IEEE Symposium on Foundations of Computer Science, 2002.
- [AP03] Dimitris Achlioptas and Yuval Peres, *The threshold for random k -SAT is $2^k(\ln 2 + o(1))$* , Proceedings of the 35th Annual ACM Symposium on Theory of Computing, 2003.
- [AS00] Dimitris Achlioptas and Gregory B. Sorkin, *Optimal myopic algorithms for random 3-SAT*, 41st Annual Symposium on Foundations of Computer Science, IEEE Comput. Soc. Press, Los Alamitos, CA, 2000, pp. 590–600.

- [AS02] D. Aldous and M. Steele, *The objective method: Probabilistic combinatorial optimization and local weak convergence*, Preprint, 2002.
- [BBC⁺01] Béla Bollobás, Christian Borgs, Jennifer T. Chayes, Jeong Han Kim, and David Bruce Wilson, *The scaling window of the 2-SAT transition*, Random Structures and Algorithms **18** (2001), no. 3, 201–256.
- [BCP97] Alberto Bertoni, Paola Campadelli, and Roberto Posenato, *An upper bound for the maximum cut mean value*, Graph-theoretic concepts in computer science (Berlin, 1997), Lecture Notes in Comput. Sci., vol. 1335, Springer, Berlin, 1997, pp. 78–84. MR 99d:68185
- [BF01] Tom Bohman and Alan M. Frieze, *Avoiding a giant component*, Random Structures and Algorithms **19** (2001), no. 1, 75–85.
- [BFU93] Andrei Z. Broder, Alan M. Frieze, and Eli Upfal, *On the satisfiability and maximum satisfiability of random 3-CNF formulas*, Proceedings of the Fourth Annual ACM-SIAM Symposium on Discrete Algorithms (Austin, TX, 1993) (New York), ACM, 1993, pp. 322–330. MR 94b:03023
- [BFW02] Tom Bohman, Alan M. Frieze, and Nicholas C. Wormald, *Avoiding a giant component II*, Unpublished manuscript, February 2002.
- [Bol88] Béla Bollobás, *Martingales, isoperimetric inequalities and random graphs*, Combinatorics (A. Hajnal, L. Lovász, and V. T. Sós, eds.), Colloq. Math. Soc. János Bolyai, no. 52, North Holland, 1988, pp. 113–139.
- [Bol98] ———, *Modern graph theory*, Springer, New York, 1998.
- [CGHS03] Don Coppersmith, David Gamarnik, Mohammad Hajiaghayi, and Gregory B. Sorkin, *Random MAX SAT, random MAX CUT, and their phase transitions*, Proceedings of the 14th Annual ACM-SIAM Symposium on Discrete Algorithms (Baltimore, MD, 2003), ACM, New York, 2003, To appear.
- [CR92] Vasěk Chvátal and Bruce Reed, *Mick gets some (the odds are on his side)*, 33th Annual Symposium on Foundations of Computer Science (Pittsburgh, PA, 1992), IEEE Comput. Soc. Press, Los Alamitos, CA, 1992, pp. 620–627.

- [Dur96] Richard Durrett, *Probability: theory and examples*, second ed., Duxbury Press, Belmont, CA, 1996. MR 98m:60001
- [FdV92] Wenceslas Fernandez de la Vega, *On random 2-SAT*, Manuscript, 1992.
- [Fri99] Ehud Friedgut, *Necessary and sufficient conditions for sharp thresholds of graph properties, and the k -SAT problem*, J. Amer. Math. Soc. **12** (1999), 1017–1054.
- [Gam02] D. Gamarnik, *Linear phase transition in random linear constraint satisfaction problem*, Preprint, 2002.
- [Goe96] Andreas Goerdt, *A threshold for unsatisfiability*, J. Comput. System Sci. **53** (1996), no. 3, 469–486.
- [GW95] Michel X. Goemans and David P. Williamson, *Improved approximation algorithms for maximum cut and satisfiability problems using semidefinite programming*, JACM **42** (1995), 1115–1145.
- [Hås97] Johan Håstad, *Some optimal inapproximability results*, STOC ’97 (El Paso, TX), ACM, New York, 1997, pp. 1–10 (electronic). MR 1715618
- [JLR00] Svante Janson, Tomasz Luczak, and Andrzej Ruciński, *Random graphs*, Wiley-Interscience Series in Discrete Mathematics and Optimization, Wiley-Interscience, New York, 2000. MR 2001k:05180
- [JS02] Svante Janson and Gregory B. Sorkin, Personal communication, 2002.
- [JSV00] Svante Janson, Yiannis C. Stamatiou, and Malvina Vamvakari, *Bounding the unsatisfiability threshold of random 3-SAT*, Random Structures Algorithms **17** (2000), no. 2, 103–116.
- [Kim] Jeong Han Kim, *Poisson cloning model for random graphs*, In preparation.
- [KKL02] A. C. Kaporis, L. M. Kirousis, and E. Lalas, *The probabilistic analysis of a greedy satisfiability algorithm*, 5-th International Symposium on the Theory and Applications of Satisfiability Testing, 2002, pp. 362–376.

- [KZ97] Howard Karloff and Uri Zwick, *A $7/8$ -approximation algorithm for MAX 3SAT?*, Proceedings of the 38th Annual IEEE Symposium on Foundations of Computer Science, Miami Beach, FL, USA, IEEE Press, 1997.
- [LLZ02] Michael Lewin, Dror Livnat, and Uri Zwick, *Improved rounding techniques for the MAX 2-SAT and MAX DI-CUT problems*, Proc. of IPCO, 2002, pp. 67–82.
- [McD89] Colin McDiarmid, *On the method of bounded differences*, London Mathematical Society Lecture Note Series, vol. 141, Cambridge University Press, 1989, pp. 148–188.
- [Spe94] Joel H. Spencer, *Ten lectures on the probabilistic method*, second ed., SIAM, 1994.
- [TSSW00] Luca Trevisan, Gregory B. Sorkin, Madhu Sudan, and David P. Williamson, *Gadgets, approximation, and linear programming*, SIAM J. Comput. **29** (2000), no. 6, 2074–2097. MR 2001j:68046
- [Wil02] David Bruce Wilson, *On the critical exponents of random k -SAT*, Random Structures and Algorithms **21** (2002), no. 2, 182–195.
- [Wor95] Nicholas C. Wormald, *Differential equations for random processes and random graphs*, Ann. Appl. Probab. **5** (1995), no. 4, 1217–1235.